

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 3, 2003, 16:55:00 ; Search time 356 Seconds
(without alignments)
15156.708 Million cell updates/sec

Title: US-09-988-863A-1
Perfect score: 2396
Sequence: 1 gtcaccaccagcgtccgycg.....ttctcaaaaaaaaaaaaaa 2396

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues
Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: N_Geneseq_101002:.*
2: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
3: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
4: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
5: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
6: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
7: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
8: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
9: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
10: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
11: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
12: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
13: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
14: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
15: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
16: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
17: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
18: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
19: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
20: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
21: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
22: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
23: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
24: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
25: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2396	100.0	2396	24	AB160244
2	443	18.5	443	24	AB194086
3	425.4	17.8	728	24	AB160246
4	299.2	12.5	611	24	AB160245
5	135.2	5.6	571	21	AA69566
6	135.2	5.6	571	21	AB160247
7	65.8	2.7	96583	21	AA62297
8	65.4	2.7	300	21	AA637035
9	57.6	2.4	1356	24	AA140805

10	57.6	2.4	1356	24	ABK96801
11	57.6	2.4	1356	24	AAD31013
12	57.6	2.4	7681	24	AAD31026
13	57.6	2.4	7693	24	AAD31023
14	57.6	2.4	7693	24	AAD31024
15	57.6	2.4	8224	24	AAD31027
16	57.6	2.4	8235	24	AAD31025
17	57.6	2.4	8400	24	AAD31029
18	57.6	2.4	13917	24	AAD31037
19	57.6	2.4	14623	24	AAD31039
20	57.6	2.4	14623	24	AAD31041
21	46	1.9	4590	22	AAH24065
22	43.6	1.8	745	21	AAC54006
23	42.8	1.8	577	22	AAD02788
24	42.8	1.8	1299	22	AAD02792
25	42.8	1.8	1299	22	AA534422
26	42.8	1.8	1763	22	AAD02791
27	41.8	1.7	164976	19	AAV21209
28	41.2	1.7	39887	22	AAK79153
29	41.2	1.7	39887	22	AAK81263
30	40.6	1.7	4590	22	AAH24065
31	40.4	1.7	116277	20	AAK20249
32	40.4	1.7	910715	20	AAK20248
33	39.2	1.6	300	23	ABV5261
34	39	1.6	3775	23	ABL12504
35	38.6	1.6	554	22	AAH10087
36	38.6	1.6	2055	22	AAH17395
37	38.6	1.6	273254	21	AAC81914
38	38.4	1.6	360	22	AAI62576
39	38.4	1.6	938	22	AAD07722
40	38	1.6	586	21	AAF08722
41	38	1.6	27681	22	AA536497
42	38	1.6	27681	22	AA536498
43	38	1.6	27681	22	AAK58843
44	37.8	1.6	1792	21	AAC59834
45	37.8	1.6	16633	24	ABN79985

ALIGNMENTS

RESULT 1	
ABL60244	standard; cDNA: 2396 BP.
ID	ABL60244
AC	ABL60244;
NC	
DT	29-JUL-2002 (first entry)
DE	Arabidopsis thaliana PMVK encoding cDNA SEQ ID NO 1.
XX	
KW	Thale cress; PMVK; phosphomevalonate kinase; plant; herbicide;
KW	growth regulator; enzyme; gene; ss.
OS	Arabidopsis thaliana.
XX	
FT	Key
FT	CDS
FT	location/Qualifiers
FT	685..2202
FT	/*tag= a
FT	/product= "PMVK"
PN	DE10057755-A1.
XX	
PD	23-MAY-2002.
XX	
PF	22-NOV-2000; 2000DE-1057755.
XX	
PR	22-NOV-2000; 2000DE-1057755.
XX	
PA	(PARB) BAYER AG.
XX	
PI	Meissner R, Lechelt-Kunze C;
XX	

S. cerevisiae pren
Yeast: phosphomeval
Operon D DNA encod
Operon A DNA encod
Operon B DNA encod
Operon E DNA encod
Operon C encoding
Operon G DNA encod
Plastid transforma
Plastid transforma
Plastid transforma
Yeast AAD9604-asso
Arabidopsis thalia
S. cerevisiae ERG
Candida albicans p
Candida albicans e
Methanococcus jan
Human immune/haema
Human immune/haema
Yeast AAD9604-asso
Borrelia burgdorfe
Human prostate exp
Drosophila melanog
Human cDNA clone (Human
Chlamydia pneumoni
Human polynucleoti
Human secreted pro
Fusarium venenatum
Human cardiovascular
Human immune/haema
Human secreted pro
Human chemically m

RESULT	6
ID	ABL60247
ABL60247	standard: DNA; 571 BP.
XX	
AC	ABL60247;
XX	
XX	29-JUL-2002 (first entry)
DE	
XX	Pinus radiata PMWK encoding DNA SEQ ID NO 5.
KW	Monterey pine; PMWK; phosphomevalonate kinase; plant; herbicide;
RW	growth regulator; ds.
OS	Pinus radiata.
XX	
PN	DE10057755-A1.
PD	
XX	23-MAY-2002.
PF	
PR	22-NOV-2000; 2000DE-1057755.
PA	22-NOV-2000; 2000DE-1057755.
PI	(FARB) BAYER AG.
DR	Weissner R, Lechelt-Kunze C;
PT	WPI: 2002-445360/48.
PT	New nucleic acid encoding plant phosphomevalonate kinase, useful for
PT	identifying modulators, potentially useful as herbicides and growth
PS	regulators -
RS	Disclosure: Page 15; 18pp; German.
CC	The invention relates to a nucleic acid (I, ABL60244) that encodes a
CC	plant phosphomevalonate kinase (PMWK, ABB77505), excluding the known
CC	fully defined partial sequences ABL60245-ABL60247. Plant PMWK; (I),
CC	constructs and host cells that contain (I) are used to identify agents
CC	that bind to and/or modulate activity of PMWK, potentially useful as
CC	herbicides and growth regulators. (I) is also used for recombinant
CC	production of PMWK. The present sequence is that of a PMWK encoding DNA
CC	excluded by the invention.
SO	
	Sequence 571 BP; 161 A; 129 C; 136 G; 145 T; 0 other;
Query Match	5.6%; Score 135.2; DB 24; Length 571;
Best Local Similarity	66.5%; Pred. No. 2.1e-29;
Matches 210; Conservative	0; Mismatches 103; Indels 3; Gaps 1.
OY	6B3 AAATGGCTGTTTGTCTTCCTCGTGGAAGTTTTGATGATGGACTGCTGTGAC 742
DB	242 ACATGGCTGTAGTTGTGTGCAGCTCCTGTGTAAGCTTTTAATAACAGGAGCTTATTATTC 301
OY	743 TCAGAGAAGCAAAATSCAGGGCTTGTGTGAGTACAATAATGCACGGTTTTACGGATTGTGA 802
DB	302 TTGAGAGAGCAAAATCCAGAGACTTGTGTGCTTACCACACAGCTGTGCTTACGCCATTGTGA 361
OY	803 AGCCAATCACAGCAAGAACGTCAGCCTGAAAAGTTGGCATGAAAAATGACAGATGTCAAAT 862
DB	362 AGCCACTCGGAGCTAGCACAGATTCACAGTAGTTGGGCAATGCGCTAATGACAGATGTGAAT 421
OY	863 TAACATACACAGAGCTCTCGAGAGAAAGCATGTATAAAGTCACAGATCATTTGACATC 922
DB	422 TAACATGCGCTTCAGCTTGCAAAAGAGGCCATTACAAAGTATCTCTGAAAGACTCTTAAGCC 481
OY	923 TTCACCTGT---GTCTGCAAGTGAATTCAAGAAACCCCTTTGTAGAGCATGCGATACAGT 979
DB	482 TGCAAAATGTGCTCTTCAACATGACATGTAATTCCTTTTGTGGAACAAGCAGTGCAAT 541
OY	980 ATGCTATAGCTGCTGC 995

	DB	542	TTCGTGTTGCACACTGC	557
	RESULT 7			
	ID	AFF2297/c		
	XX	AFF22297 standard; DNA;	96583 BP.	
	AC			
	XX	AFF22297;		
	DT	20-MAR-2001 (first entry)		
	DE	BAC containing repeats from centromeres 1-4 #20.		
	-XX	Centromere; microsome; vector; ds.		
	KW			
	OS	Arabidopsis thaliana.		
	PN	WO200055325-A2.		
	XX			
	PD	21-SEP-2000.		
	PB	17-MAR-2000; 2000WO-US07392.		
	PR	18-MAR-1999; 99US-0125219.		
	PR	01-APR-1999; 99US-0127409.		
	PR	18-MAY-1999; 99US-0134770.		
	PR	13-SEP-1999; 99US-0153584.		
	PR	17-SEP-1999; 99US-0154603.		
	XX			
	PA	(UYCH-) UNIV CHICAGO.		
	XX			
	PI	Preuss D, Copenhagen G, Keith K;		
	DR	WPI; 2000-587529/55.		
	PT	Recombinant DNA construct comprising a plant centromere, useful for producing stably inherited microsomes which can serve as vectors for the construction of transgenic plant and animal cells		
	PS	Claim 102; Page 716-738; 1449pp; English.		
	CC	The present invention relates to a recombinant DNA construct of a plant (Arabidopsis thaliana) centromere. The constructs are useful for producing stably inherited microsomes which can serve as vectors for the construction of transgenic plant and animal cells expressing selected proteins such as hormones, enzymes, interleukins, clotting factors, cytokines, antibodies, and growth factors.		
	SO	Sequence 96583 BP; 29910 A; 18944 C; 18341 G; 29388 T; 0 other;		
	Query Match	2.7%; Score 65.8; DB 21; Length 96583;		
	Best Local Similarity	61.3%; Pred. No. 1.5e-07;		
	Matches 106; Conservative	0; Mismatches 67; Indels 0; Gaps 0;		
OY	100	AAACCCTATAACCGTGGTGACAGAGTGCCGAACCGAACGCCGTAACCGAATC	159	
Db	20279	AAACCCGAAGTAATGATGATCACTGATTNCCAACAACGACCAACGACATCATGATTG	20220	
OY	160	CCTCAAAGAATTTGCCGATCGGTTTGCTACACTGTCCAAAACCTCGGTGCCGAGACCGAAA	219	
Db	20219	CCCATAATTAATTTGGCGGTGGTTGGTTTAACTCTTCCGGCTAAAGCCGANAACCGAAA	20160	
OY	220	CTGTCGGTTTTTGGGTTCCGCGGTCTCTCGAAGTCCGAGGCCCA	272	
Db	20159	ATTTCGGGTGATCGTCCGCGGTTCGATGCGTTTAAATCCGAACCTCCAGGCCCA	20107	
	RESULT 8			
	ID	AAC37035		
	XX	AAC37035 standard; DNA; 300 BP.		
	AC			
	XX	AAC37035;		

DT 17-OCT-2000 (first entry)
XX Arabidopsis thaliana DNA fragment SEQ ID NO: 15946.
XX Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX Arabidopsis thaliana.
OS
PN EPI033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132853.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.

PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0144806.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 06-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 18-AUG-1999; 99US-0149427.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.

27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159285.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 2.7%; Score 65.4; DB 21; Length 300;
Best Local Similarity 98.5%; Pred. No. 7.7e-09;
Matches 66; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2096 TTTTGGCTTTGTGGTAGAGATCCACATGGGCTTGGCTAGAAAGTGTATCCAC 2155
DB 234 TTGGGCGCTTGTGGTAGAGATCCACATGGGCTTGGCTAGAAAGTGTATCCAC 293
QY 2156 GAACCAC 2162
DB 294 GAACCAC 300

RESULT 9

ID AAL40805 standard; DNA; 1356 BP.

AC AAL40805;

DT 03-OCT-2002 (first entry)

DE Nucleic acid relating to the production of prenyl alcohol SEQ ID NO 29.

XX Prenyl alcohol; mutated cell; squalene synthase gene; geometric isomer;
KW Industrial synthesis; isoprenoid-terpenoid compound; ds.
XX Saccharomyces cerevisiae.
OS WO200253747-A1.
PN 11-JUL-2002.
PD 20-DEC-2001; 2001WO-JP11215.
PE 28-DEC-2000; 2000JP-0401701.
PR 28-DEC-2000; 2000JP-0403067.
PR 18-SEP-2001; 2001JP-0282978.
XX (TOYT) TOYOTA JIDOSHA KK.
PA Onto C, Obata S;
PI WPI; 2002-548086/58.
DR Production of prenyl alcohols by culturing translationally-active
XX mutated cells with reduced squalene synthase gene to express less
PT transcriptional product -
PS Disclosure; Page 237-238; 262pp; Japanese.
XX The invention relates to a method for producing prenyl alcohol comprising
CC culturing mutated cells, having been mutated so as to reduce the amount
CC of transcriptional product of squalene synthase gene transcriptional
CC activity, and then collecting prenyl alcohol from the culture medium. The
CC method is for the production of prenyl alcohols, which is for use in
CC industrial synthesis of isoprenoid-terpenoid compounds particularly
CC physiologically-active prenyl alcohol geometric isomers. This
CC polynucleotide sequence represents a nucleic acid sequence relating to
CC the method for producing prenyl alcohol comprising culturing mutated
CC cells.
XX
SQ Sequence 1356 BP; 427 A; 236 C; 304 G; 389 T; 0 other:

Query Match 2.4%; Score 57.6; DB 24; Length 1356;
Best Local Similarity 50.7%; Pred. No. 3.7e-06;
Matches 138; Conservative 0; Mismatches 134; Indels 0; Gaps 0;

QY 1337 ATCTAGATGTTATCCATATGATGATGACACAAAGCTCATGTTGTCACAAAGGAGTCTG 1396
DB 551 ATAGAGAGTTATTCATATTTAGCACAAAGTGTCTCATTTGCAAGCTCAGGTAATAATG 610
QY 1397 GAATGGGTTGATGTCAGCTGTCTGTATGGAAGTCACAGCTTATGTTGGCTTCTCTC 1456
DB 611 GAACGGGTTGATGATGACGGCGCAGCATATGATGATATGATGATGATGATGATGATG 670
QY 1457 CAGAAGTCTTCTCATTTGCTCAGGTTGCGATGATGATGATGATGATGATGATGATG 1516
DB 671 CCGCATTTATCTCTAATTTGCGCATATGATGATGATGATGATGATGATGATGATGATG 730
QY 1517 GTACATTTTGAAGGAAATGGAACATTAAGAGAACTGATCTTTTACACCACTGA 1576
DB 731 ATTTGGTTGATGAGAAAGATGGAATATTTAGCAATTAAGTAACCAATTTACTCTGGGAT 790
QY 1577 TGAATCTTTTCTTGAGAGAACCTGGAAGTGT 1608
DB 791 TAACTTATGATGGCGCATATTAAGATGT 822

RESULT 10

ID ABR96801

ID ABR96801 standard; DNA; 1356 BP.

AC ABR96801;

DT 24-SEP-2002 (first entry)

DE S. cerevisiae prenyl diphosphate synthase gene #20.

XX
XX
KW prenyl alcohol; prenyl diphosphate synthase; geranylgeraniol;
KM hydroxymethylglutaryl-CoA reductase; farnesylgeraniol; gene; ss;
KW isopentenyl diphosphate delta-isomerase; mevalonate kinase;
RV mevalonate CoA acetyltransferase; isoprenoid-terpenoid compound.
XX
XX Saccharomyces cerevisiae.
XX
XX MO200253746-A1.
PN
PN 11-JUL-2002.
PD
XX 20-DEC-2001; 2001WO-JP11214.
PE
PR 28-DEC-2000; 2000JP-0403067.
PR
XX (TOYT) TOYOTA JIDOSHA KK.
PA
PI Ohto C, Obata S, Muramatsu M, Nishi K, Totsuka K;
PI WPI; 2002-537944/57.
DR
XX
XX
PT Production of prenyl alcohols by culturing a transformant transferred
PT with e.g. prenyl diphosphate synthase gene, for use in industrial
PT synthesis of e.g. physiologically-active isoprenoid-terpenoid compounds
PT
XX
XX
PS Example 17; Page 269-270; 33pp; Japanese.
XX
XX The invention relates to a process for producing a prenyl alcohol
CC comprising: (A) construction of a recombinant by transferring an
CC expression recombinant DNA or a DNA for genome integration into a host
CC which contains prenyl diphosphate synthase gene or its variant; and
CC (B) collecting product from the culture medium. Also described is: (1) a
CC method for producing prenyl alcohol in which the expression
CC recombinant DNA contains: (a) a hydroxymethylglutaryl-CoA reductase
CC gene or its variant; or (b) an isopentenyl diphosphate delta-isomerase
CC gene; (2) a process for producing geranylgeraniol, comprising:
CC (a) constructing a recombinant by transferring an expression
CC recombinant DNA or a DNA for genome integration into a host which
CC contains hydroxymethylglutaryl-CoA reductase gene or its variant; and
CC (b) isolating the product; (3) a process for producing farnesylgeraniol
CC in which the expression recombinant DNA also contains an isopentenyl
CC diphosphate delta-isomerase gene, mevalonate CoA acetyltransferase gene,
CC hydroxymethylglutaryl-CoA synthase gene, mevalonate kinase gene, or
CC mevalonate diphosphate decarboxylase gene. The methods are used for
CC the production of prenyl alcohols, particularly for use in industrial
CC synthesis of isoprenoid-terpenoid compounds e.g. physiologically-active
CC prenyl alcohols including geometric isomers. ABX96780-ABX96897
CC represent prenyl diphosphate synthase genes and related PCR primers used
CC in the methods of the invention.
XX
XX

SO Sequence 1356 BP; 427 A; 236 G; 304 G; 389 T; 0 other;

Query Match 2.4%; Score 57.6; DB 24; Length 1356;
Best Local Similarity 50.7%; Pred. No. 3.7e-06;
Matches 138; Conservative 0; Mismatches 134; Indels 0; Gaps

0Y 1337 ATCTAGATGTTATCCATATGATAGCACAAACGTCCTATTGTCTTGCAACAAGGAGGTGC 1396
DB 551 ATAGCAAGATGTTATCCAAATTTATACCAACAACTGCTCTTGTCAACGACAGGATTAATTG 610
0Y 1397 GAATGGATTTGATGTCAGCGTCGCTCTCATGGAAGTCAAGCGTTATAGTTCGCTCTCTC 1456
DB 611 GAAGCGGTTTGAATGATGACCGCGGCACCATATGGATCTATACATATAGAAATTTCCAC 670
0Y 1457 CAGAAGCTTGCTCATTTTGCTCAGGTTGCTACAGTACAGGTGCGCATTAAGAAGATTATTG 1516
DB 671 CCGCATTAATCTTAATTTGGCCAAATATTGGAAGTGTCTACTTACGCGGAGTAAACTGGCCG 730
0Y 1517 GTACATATTTTGAAGGAAATGGGACATATTAAGACACTGAGTTCCTTTTACCAACCACTGA 1576

Dd		791	ATTTCGTTGATGAGAAACACTGGAAATATTACATAAAGTAACCATTTACCTTGCGGAT	790
Oy	1577	TGAATCTTTTCCCTTGAGAACCTGGAAGTGT	1608	
Dd	791	TAACTTTATGCATGGCGCATATTATGAATGCT	822	
	RESULT 11			
ID	AAD31013			
XX	AAD31013 standard; DNA; 1356 BP.			
AC	AAD31013;			
XX				
Dt	31-MAY-2002 (first entry)			
XX				
DE	Yeast phosphomevalonate kinase (ERG8)orf.			
XX				
KW	Mevlonate: Isopentenyl diphosphate; herbicide resistance; DMAPP; IPP;			
RW	Isoprenoid; dimethylallyl diphosphate; antibiotic resistance; enzyme;			
KW	transgenic plant; yeast; phosphomevalonate kinase; ERG8; PMK; ds.			
XX				
OS	Saccharomyces cerevisiae.			
XX	WO200210398-A2.			
PN				
XX	07-FEB-2002.			
PD				
XX				
PE	31-JUL-2001; 2001MO-US24037.			
PR				
XX	31-JUL-2000; 2000US-221703P.			
PA	(HAHN/) HAHN F M.			
PA	(KUEH/) KUEHNLE A R.			
XX				
Pf	Hahn FM, Kuehnle AR:			
PI				
XX				
DR	WPI: 2002-217122/27.			
XX				
Pt	use of specific genes of mevalonate and isoprenoid biosynthetic			
Pt	pathways, for providing a cell with herbicide or antibiotic resistance,			
Pt	and for providing transformed cells having increased isoprenoid			
Pt	production -			
PT				
PS	Disclosure: Page 116-117; 193pp; English.			
XX				
CC	The invention relates to the use of specific genes of the mevalonate			
CC	and isoprenoid biosynthetic pathways and inactive gene sites			
CC	(pseudogene). Genes of the invention are used to enhance biosynthesis			
CC	of isopentenyl diphosphate (IPP), dimethylallyl diphosphate (DMAPP)			
CC	and isoprenoid pathway derived products in the plastids of transgenic			
CC	plants and microalgae, for producing herbicide or antibiotic resistant			
CC	transgenic plants and microalgae, for providing transformed cells with			
CC	increased isoprenoid production compared to non-transformed cells, and			
CC	for providing a cell with an inserted polynucleotide sequence encoding			
CC	one or more products of interest. The present sequence is yeast			
CC	phosphomevalonate kinase (PMK; ERG8) EC 2.7.4.2. encoding orf.			
XX				
SO	Sequence 1356 BP; 427 A; 236 C; 304 G; 389 T; 0 other:			
	Query Match 2.4%; Score 57.6; DB 24; Length 1356;			
	Best Local Similarity 50.7%; Pred.No.3.7e-06;			
	Matches 138; Conservative 0; Mismatches 134; Indels 0; Gaps 0;			
Oy	1337 ATCTAGATGTTTCATATGATAGCAACAACGCTCATTGCTGCACAGAAGGANGTCG	1396		
Dd	551 ATAGAGAGATTATTCATATTTAGACAAAGTTCATTGCTCAAGCTCAGGGTAAAATG	610		
Oy	1397 GAAGGAGGTTTGATGTCAGCTGCTGCTATATGAGAAAGTCAGCGTATTCTCCCTTC	1456		
Dd	611 GAAGGCGGTTTATGTAAGCGCGCGCACGATATGATATCAGATATAGAACATTCGCCAC	670		
Oy	1457 CAGAGCTCTGTGCATTTGCTCAGGTTGACAGTAACAGGCTGCCATTAAATGAAGTATTG	1516		

Db 671 CCGCATTAATCTCTAATTTGGCAGATATTTGAGAGTGTACTTACGCGAGTAACTGGCGC 730
 Oy 1517 GTACAATTTTGAAGGAAATGGACATATAGAGAACTGAGTCTCTTACCACACTGA 1576
 Db 731 ATTTGGTTGATGAGAAAGACTGGAATATTTACGATTTAAAGTAACCATTTTACCTTCGGGAT 790
 Oy 1577 TGAATCTTTTCTTGAGAACCTGGAAGTGT 1608
 Db 791 TAACCTTATGATGGCGGATTTAAGATGT 822

RESULT 12
 AAD31026 standard; DNA: 7681 BP.
 AAD31026:
 31-MAY-2002 (first entry)
 Operon D DNA encoding mevalonate pathway.
 Mevalonate; isopentenyl diphosphate; herbicide resistance; DMAPP; IPP;
 isoprenoid; dimethylallyl diphosphate; antibiotic resistance; enzyme;
 transgenic plant; yeast; phosphomevalonate kinase; HMGR; HMGs; AACT;
 mevalonate kinase; mevalonate diphosphate decarboxylase; HMG-CoA; MDD;
 acetocetyl thiolase; 3-hydroxy-3-methylglutaryl-coenzyme A synthase;
 MK; PMK; HMG-CoA reductase; ds.
 Chimeric - Saccharomyces cerevisiae.
 Chimeric - Arabidopsis thaliana.
 Chimeric - Streptomyces sp.
 WO200210398-A2.
 07-FEB-2002.
 31-JUL-2001; 2001WO-US24037.
 31-JUL-2000; 2000US-221703P.
 (HAHN/) HAHN F. M.
 (KUEH/) KUEHNLE A. R.
 Hahn FM, Kuehnle AR;
 WPI: 2002-217122/27.
 Use of specific genes of mevalonate and isoprenoid biosynthetic
 pathways, for providing a cell with herbicide or antibiotic resistance,
 and for providing transformed cells having increased isoprenoid
 production -
 Claim 77: Page 141-145; 193pp: English.
 The invention relates to the use of specific genes of the mevalonate
 and isoprenoid biosynthetic pathways and inactive gene sites
 (pseudogene). Genes of the invention are used to enhance biosynthesis
 of isopentenyl diphosphate (IPP), dimethylallyl diphosphate (DMAPP)
 and isoprenoid pathway derived products in the plastids of transgenic
 plants and microalgae, for producing herbicide or antibiotic resistant
 transgenic plants and microalgae, for providing transformed cells with
 increased isoprenoid production compared to non-transformed cells, and
 for providing a cell with an inserted polynucleotide sequence encoding
 one or more products of interest. The present sequence is operon D DNA
 encoding the entire mevalonate pathway. This operon contains
 S. cerevisiae orf encoding phosphomevalonate kinase (PMK), mevalonate
 kinase (MWK), mevalonate diphosphate decarboxylase (MDD), acetocetyl
 thiolase (AACT) and A. thaliana orf encoding 3-hydroxy-3-methylglutaryl-
 coenzyme A (HMG-CoA) synthase (HMGs) and Streptomyces sp. CL190 orf
 encoding HMG-CoA reductase (HMGRt).
 Sequence 7681 BP; 2127 A; 1691 C; 1837 G; 2026 T; 0 other;

Query Match 2.4%; Score 57.6; DB 24; Length 7681;
 Best Local Similarity 50.7%; Pred. No. 9.7e-06;
 Matches 138; Conservative 0; Mismatches 134; Indels 0; Caps 0;

Oy 1337 ATCTAGATGTTATTCATATGATACACAAAGCTCTCTTGTCTGCACAAAGGAAAGTGC 1396
 Db 585 ATAGACAGTTATTCATTAATTTAGCAAGTGTCTCATTTGTCAGGTAATAATG 644
 Oy 1397 GAATGGGTTTGGATGTCAGCTGCTCTGTCATGAACTACAGCTTATGTCCTCTCTC 1456
 Db 645 GAACGGGTTTGGATGTCAGCGCGGACACATATGATCTATCAGATATAGAAATTCAC 704
 Oy 1457 CAGAGTCTTGTCTCATTTGCTCAGGTGACAGTACAGGTCTGCCATTAAAGAACTATTG 1516
 Db 705 CCGCATTAATCTCTAATTTGGCAGATATTTGAGAGTGTACTTACGCGAGTAACTGGCGC 764
 Oy 1517 GTACAATTTTGAAGGAAATGGACATATAGAGAACTGAGTCTCTTACCACACTGA 1576
 Db 765 ATTTGGTTGATGAGAAAGACTGGAATATTTACGATTTAAAGTAACCATTTTACCTTCGGAT 824
 Oy 1577 TGAATCTTTTCTTGAGAACCTGGAAGTGT 1608
 Db 825 TAACCTTATGATGGCGGATTTAAGATGT 856

RESULT 13
 AAD31023 standard; DNA: 7693 BP.
 AAD31023:
 31-MAY-2002 (first entry)
 Operon A DNA encoding mevalonate pathway.
 Mevalonate; isopentenyl diphosphate; herbicide resistance; DMAPP; IPP;
 isoprenoid; dimethylallyl diphosphate; antibiotic resistance; enzyme;
 transgenic plant; yeast; phosphomevalonate kinase; HMGR; HMGs; AACT;
 mevalonate kinase; mevalonate diphosphate decarboxylase; HMG-CoA; MDD;
 acetocetyl thiolase; 3-hydroxy-3-methylglutaryl-coenzyme A synthase;
 MK; PMK; HMG-CoA reductase; ds.
 Chimeric - Saccharomyces cerevisiae.
 Chimeric - Arabidopsis thaliana.
 Chimeric - Streptomyces sp.
 WO200210398-A2.
 07-FEB-2002.
 31-JUL-2001; 2001WO-US24037.
 31-JUL-2000; 2000US-221703P.
 (HAHN/) HAHN F. M.
 (KUEH/) KUEHNLE A. R.
 Hahn FM, Kuehnle AR;
 WPI: 2002-217122/27.
 Use of specific genes of mevalonate and isoprenoid biosynthetic
 pathways, for providing a cell with herbicide or antibiotic resistance,
 and for providing transformed cells having increased isoprenoid
 production -
 Claim 77: Page 127-131; 193pp: English.
 The invention relates to the use of specific genes of the mevalonate
 and isoprenoid biosynthetic pathways and inactive gene sites
 (pseudogene). Genes of the invention are used to enhance biosynthesis
 of isopentenyl diphosphate (IPP), dimethylallyl diphosphate (DMAPP)
 and isoprenoid pathway derived products in the plastids of transgenic

CC plants and microalgae, for producing herbicide or antibiotic resistant
 CC transgenic plants and microalgae, for providing transformed cells with
 CC increased isoprenoid production compared to non-transformed cells, and
 CC for providing a cell with an inserted polynucleotide sequence encoding
 CC one or more products of interest. The present sequence is operon A DNA
 CC encoding the entire mevalonate pathway. This operon contains
 CC 5. cerevisiae orfs encoding phosphomevalonate kinase (PMK), mevalonate
 CC kinase (MVK), mevalonate diphosphate decarboxylase (MDD), acetocetyl
 CC thiolase (ACT) and A. thaliana orfs encoding 3-hydroxy-3-methylglutaryl-
 CC coenzyme A (HMG-CoA) synthase (HMGs) and HMG-CoA reductase (HMGrt).
 XX

SQ Sequence 7693 BP; 2212 A; 1526 C; 1794 G; 2161 T; 0 other;

Query Match 2.4%; Score 57.6; DB 24; Length 7693;

Best Local Similarity 50.7%; Pred. No. 9.8e-06;
 Matches 138; Conservative 0; Mismatches 134; Indels 0; Gaps 0;

OY 1337 ATCTAGATGTTATCATATGATGACAAACGCTCATTTGTTCCACAAAGGAAGTCG 1396

DB 4338 ATAGAGAGTTATTCATATTTAGCAAGTGTCTCATTTGCAAGCTCAGGTTAAATG 4397

OY 1397 GAAGTGGTTGATGTCAGCTGCTGCTATGAGAGTCAGGTTATGTCCTCTCTC 1456

DB 4398 GAAGCGGTTGATGTCAGCGGCGGCGATATGATCTATGATAGATAGAAATTCAC 4457

OY 1457 CAGAGCTTGTCTCATTTGCTCAGTTCAGTACAGCTGCTGCATTAATGAATATG 1516

DB 4458 CCGCATTTATCTCTAATTTGCCAGATATTGGAAGTCTACTACGCACTAAACCTGGCC 4517

OY 1517 GTACAATTTTGAAGGAAATGGACAATAGAGAAGTGAAGTCTCTTTACCAACACATGA 1576

DB 4518 ATTTGGTTGATGAGAAAGACATGGAATATTCGATTAAAGTAACATTACCTCGGAT 4577

OY 1577 TGAATCTTTTCTTGAGAACCTGGAAGTGT 1608

DB 4578 TAACTTTATGATGGCGATATTAAAGATGTT 4609

RESULT 14
 AAD31024
 ID AAD31024 standard; DNA; 7695 BP.

AC AAD31024;

DT 31-MAY-2002 (first entry)

XX Operon B DNA encoding mevalonate pathway.

KW Mevalonate; isopentenyl diphosphate; herbicide resistance; DMAPP; IPP;
 KW isoprenoid; dimethylallyl diphosphate; antibiotic resistance; enzyme;
 KW transgenic plant; Yeast; phosphomevalonate kinase; HMGrt; HMGs; ACT;
 KW mevalonate kinase; mevalonate diphosphate decarboxylase; HMG-CoA; MDD;
 KW acetocetyl thiolase; 3-hydroxy-3-methylglutaryl-coenzyme A synthase;
 KW MVK; PMK; HMG-CoA reductase; ds.

XX Chimeric - Saccharomyces cerevisiae.

OS Chimeric - Arabidopsis thaliana.

XX WO200210398-A2.

PD 07-FEB-2002.

XX 31-JUL-2001; 2001WO-US24037.

PR 31-JUL-2000; 2000US-221703P.

PA (HAHN/) HAHN F M.

PA (KUEH/) KUEHNLE A R.

PI Hahn FM, Kuehnle AR;

DR WPI: 2002-217122/27.

XX

PT Use of specific genes of mevalonate and isoprenoid biosynthetic
 PT pathways, for providing a cell with herbicide or antibiotic resistance,
 PT and for providing transformed cells having increased isoprenoid
 PT production

PS Claim 77; Page 132-136; 193pp; English.

CC The invention relates to the use of specific genes of the mevalonate
 CC and isoprenoid biosynthetic pathways and inactive gene sites
 CC (pseudogene). Genes of the invention are used to enhance biosynthesis
 CC of isopentenyl diphosphate (IPP), dimethylallyl diphosphate (DMAPP)
 CC and isoprenoid pathway derived products in the plastids of transgenic
 CC plants and microalgae, for producing herbicide or antibiotic resistant
 CC transgenic plants and microalgae, for providing transformed cells with
 CC increased isoprenoid production compared to non-transformed cells, and
 CC for providing a cell with an inserted polynucleotide sequence encoding
 CC one or more products of interest. The present sequence is operon B DNA
 CC encoding the entire mevalonate pathway. This operon contains
 CC 5. cerevisiae orfs encoding phosphomevalonate kinase (PMK), mevalonate
 CC kinase (MVK), mevalonate diphosphate decarboxylase (MDD), acetocetyl
 CC thiolase (ACT) and A. thaliana orfs encoding 3-hydroxy-3-methylglutaryl-
 CC coenzyme A (HMG-CoA) synthase (HMGs) and HMG-CoA reductase (HMGrt).
 XX

SQ Sequence 7695 BP; 2214 A; 1524 C; 1793 G; 2164 T; 0 other;

Query Match 2.4%; Score 57.6; DB 24; Length 7695;

Best Local Similarity 50.7%; Pred. No. 9.8e-06;
 Matches 138; Conservative 0; Mismatches 134; Indels 0; Gaps 0;

OY 1337 ATCTAGATGTTATCATATGATGACAAACGCTCATTTGTTCCACAAAGGAAGTCG 1396

DB 570 ATAGAGAGTTATTCATATTTAGCAAGTGTCTCATTTGCAAGCTCAGGTTAAATG 629

OY 1397 GAAGTGGTTTATGTCAGCTGCTGCTATGAGAGTGAAGTATGTCCTCTCTC 1456

DB 630 GAAGCGGTTTATGTCAGCGGCGGCGATATGATCTATGATAGATAGAAATTCAC 689

OY 1457 CAGAGCTTGTCTCATTTGCTCAGTTCAGTACAGCTGCTGCATTAATGAATATG 1516

DB 690 CCGCATTTATCTCTAATTTGCCAGATATTGGAAGTCTACTACGCACTAAACCTGGCC 749

OY 1517 GTACAATTTTGAAGGAAATGGACAATAGAGAAGTGAAGTCTCTTTACCAACACATGA 1576

DB 750 ATTTGGTTGATGAGAAAGACATGGAATATTCGATTAAAGTAACATTACCTCGGAT 809

OY 1577 TGAATCTTTTCTTGAGAACCTGGAAGTGT 1608

DB 810 TAACTTTATGATGGCGATATTAAAGATGTT 841

RESULT 15
 AAD31027
 ID AAD31027 standard; DNA; 8224 BP.

AC AAD31027;

DT 31-MAY-2002 (first entry)

XX Operon E DNA encoding mevalonate pathway and IPP isomerase.

KW Mevalonate; isopentenyl diphosphate; herbicide resistance; DMAPP; IPP;
 KW isoprenoid; dimethylallyl diphosphate; antibiotic resistance; enzyme;
 KW transgenic plant; Yeast; phosphomevalonate kinase; HMGrt; HMGs; ACT;
 KW mevalonate kinase; mevalonate diphosphate decarboxylase; HMG-CoA; MDD;
 KW acetocetyl thiolase; 3-hydroxy-3-methylglutaryl-coenzyme A synthase;
 KW MVK; PMK; HMG-CoA reductase; IPP isomerase; IPP; ds.

XX Chimeric - Saccharomyces cerevisiae.

OS Chimeric - Arabidopsis thaliana.

OS Chimeric - Streptomyces sp.

XX Chimeric - Rhodobacter capsulatus.

XX WO200210398-A2.

XX 07-FEB-2002.
PD
XX
XX 31-JUL-2001; 2001MO-US24037.
PE
XX
XX 31-JUL-2000; 2000US-221703P.
PR
XX
XX (HAHN/) HAHN F. M.
PA (KUEH/) KUEHNLE A. R.
XX
PI Hahn FM, Kuehnle AR;
XX
DR WPI; 2002-217122/27.
XX
XX
PT Use of specific genes of mevalonate and isoprenoid biosynthetic
PT pathways, for providing a cell with herbicide or antibiotic resistance,
PT and for providing transformed cells having increased isoprenoid
PT production
XX
XX
PS Claim 77; Page 145-149; 193pp; English.
XX
XX
CC The invention relates to the use of specific genes of the mevalonate
CC and isoprenoid biosynthetic pathways and inactive gene sites
CC (pseudogene). Genes of the invention are used to enhance biosynthesis
CC of isopentenyl diphosphate (IPP), dimethylallyl diphosphate (DMAPP)
CC and isoprenoid pathway derived products in the plastids of transgenic
CC plants and microalgae, for producing herbicide or antibiotic resistant
CC transgenic plants and microalgae, for providing transformed cells with
CC increased isoprenoid production compared to non-transformed cells, and
CC for providing a cell with an inserted polynucleotide sequence encoding
CC one or more products of interest. The present sequence is operon B DNA
CC encoding the entire mevalonate pathway. This operon contains
CC S. cerevisiae orf5 encoding phosphomevalonate kinase (PMK), mevalonate
CC kinase (MKK), mevalonate diphosphate decarboxylase (MDD), acetoacetyl
CC thiolase (AACT) and A. thaliana orf5 encoding 3-hydroxy-3-methylglutaryl-
CC coenzyme A (HMG-CoA) synthase (HMGs); Streptomyces sp. CL130 orf
CC encoding HMG-CoA reductase (HMGR) and R. capsulatus DNA encoding IPP
CC isomerase (IPPI).
XX
SQ Sequence 8224 BP; 2228 A; 1847 C; 2016 G; 2133 T; 0 other;
XX
XX
Query Match 2.48; Score 57.6; DB 24; Length 8224;
Best Local Similarity 50.78; Pred. No. 1e-05;
Matches 138; Conservative 0; Mismatches 134; Indels 0; Gaps 0;
XX
QY 1337 ATCTAGATGTTATCCATATGATATGACACAAACGCTCATTTGTCACACAGGAAAGTGC 1396
DB 585 ATAGAGAGATTATTCATATATTAGCACAAAGTTGCTCATTTGTCACAGGTAATAATTG 644
QY 1397 GAATGGGTTGATGTCAGCTGCTGCTATGGAAGTCAGCTTATGCTTCTCTC 1456
DB 645 GAACGGGTTGATGTCAGCTGCTGCTATGGAAGTCAGCTTATGGAAGATTCCAC 704
QY 1457 CAGAAGCTTGTCTCATTTGCTCAGCTTGCAGTAAACAGTCCCATTTAATGAAGTTATTG 1516
DB 705 CCGCATTAATCTCTAATTTTGCACATATGGAAGTCTACTTACGGCAGTAACTGGCGC 764
QY 1517 GTACAATTTTGAAGAAATGAGACATAAGAGAACTGAGTCTTACACACACATGA 1576
DB 765 ATTGGTTGATGAGAAAGACTGGAATATTTACGATTAAAGTAACATTACCTCGGAT 824
QY 1577 TGAATCTTTTCTTGAGAGAACTGGAAGTGT 1608
DB 825 TAACCTTATGATGGCGCATTTAAGATGCT 856

Search completed: May 3, 2003, 18:09:32
Job time : 597 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 3, 2003, 16:56:55 ; Search time 4228 Seconds

(without alignments)
16492.508 Million cell updates/sec

Title: US-09-988-863A-1

Perfect score: 2396

Sequence: 1 gtcgaccacagctgcgycg.....ttctcaaaaaaaaaaaaaa 2396

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:*

1: gb_da:*

2: gb_hcg:*

3: gb_in:*

4: gb_om:*

5: gb_ov:*

6: gb_pat:*

7: gb_ph:*

8: gb_pl:*

9: gb_pr:*

10: gb_ro:*

11: gb_sts:*

12: gb_sy:*

13: gb_un:*

14: gb_vl:*

15: em_ba:*

16: em_fun:*

17: em_hum:*

18: em_in:*

19: em_mu:*

20: em_om:*

21: em_or:*

22: em_ov:*

23: em_pat:*

24: em_ph:*

25: em_pl:*

26: em_ro:*

27: em_sts:*

28: em_un:*

29: em_vl:*

30: em_hcg_hum:*

31: em_hcg_inv:*

32: em_hcg_other:*

33: em_hcg_mus:*

34: em_hcg_plo:*

35: em_hcg_rtd:*

36: em_hcg_mam:*

37: em_hcg_vrt:*

38: em_sy:*

39: em_hgtgo_hum:*

40: em_hgtgo_mus:*

41: em_hgtgo_other:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2396	100.0	2396	6	AX441242 Sequence
2	841.2	35.1	1785	8	AF429385 Hevea bra
3	759	31.7	119420	8	AC079041 Arabidops
4	425.4	17.8	728	6	AX441245 Sequence
5	299.2	12.5	611	6	AX441244 Sequence
6	227	9.5	40548	8	AB011480 Arabidops
7	135.2	5.6	571	6	AX441246 Sequence
8	123.2	5.1	2004	6	AX461344 Sequence
9	123.2	5.1	8840	8	AC004705 Arabidops
10	123.2	5.1	92524	8	AC005398 Arabidops
11	108.2	4.5	114041	8	AC005957 Arabidops
12	93.2	3.9	81877	8	T32B20 Arabidops
13	91.4	3.8	54386	8	AP002068 Arabidops
14	87.4	3.6	96183	8	T518 Arabidops
15	86	3.6	108061	8	AC009526 Arabidops
16	86	3.6	131692	8	AC006423 Arabidops
17	83.8	3.5	82214	8	AC023913 Arabidops
18	83.8	3.5	117585	8	AC023279 Genomic s
19	81	3.4	61712	8	AB017061 Arabidops
20	81	3.4	126253	8	AC019018 Arabidops
21	80.8	3.4	110565	8	AC009243 Genomic s
22	80.2	3.3	54252	8	AB017060 Arabidops
23	78.8	3.3	65839	8	AC005916 Arabidops
24	78.4	3.3	81414	8	AB024037 Arabidops
25	78.4	3.3	110157	8	AF058825 Arabidops
26	77	3.2	85020	8	AB018113 Arabidops
27	76.6	3.2	2000	6	AX461275 Sequence
28	76.6	3.2	95870	8	AC007396 Genomic s
29	76.6	3.2	132522	8	AB062089 Arabidops
30	73.6	3.1	42947	8	SPAC343 Arabidops
31	72.6	3.0	105733	8	AC012561 Arabidops
32	70.8	3.0	74475	8	T7M24 Arabidops
33	70.8	3.0	196296	8	ATCHRIV10 Arabidops
34	70.2	2.9	47412	6	AX059543 Sequence
35	70.2	2.9	122529	8	T1J24 Arabidops
36	70.2	2.9	199987	8	ATCHRIV15 Arabidops
37	70	2.9	92260	8	AC006194 Arabidops
38	69.6	2.9	82098	8	AC006136 Arabidops
39	69.4	2.9	112862	8	AC025781 Arabidops
40	65.8	2.7	42112	6	AX059497 Sequence
41	65.8	2.7	96583	8	ATAC009992 Arabidops
42	63.6	2.7	90627	8	AP000411 Arabidops
43	61.6	2.6	92049	8	AC004482 Arabidops
44	60	2.5	115641	8	FRK4 Arabidops
45	59.8	2.5	81341	8	AC007213 Arabidops

ALIGNMENTS

RESULT 1

AX441242

LOCUS AX441242 2396 bp DNA linear PAT 02-JUL-2002

DEFINITION Sequence 1 from Patent EP1209236.

ACCESSION AX441242

VERSION AX441242.1 GI:21690239

KEYWORDS

SOURCE

ORGANISM

thale cress.

Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE

AUTHORS

Meissner, R. and Lechelt-Kunze, C.

TITLE

Phosphomevalonate kinases from plants

Db	1981	GCTGAGAGAGTTCTACTGTGCTGCTGCTGAGCTGGTGGATTGGATGCCATTATTGCA	2040
QY	2041	ATCACTTTAGGGGATTTCCGGCACCAACTGACCCAGCAGTGGAGTTGGCACANTGTTTGG	2100
Db	2041	ATCACTTTAGGGGANTCCGGCACCAACTGACCCAGGAGTGGAGTTGCCACANTGTTTGG	2100
QY	2101	GCCTTGTTGGTGAGAGAAATCCACATGCGGCTTCCATAGAAAGTGGATCCAGCAACC	2160
Db	2101	GCCTTGTTGGTGAGAGAAATCCACATGCGGCTTCCATAGAAAGTGGATCCAGCAACC	2160
QY	2161	ACATGTATTACTTCAAGCGTTTCATCAATTCACCTTGAGTAAACAACATGTTTCAGTGT	2220
Db	2161	ACATGTATTACTTCAAGCGTTTCATCAATTCACCTTGAGTAAACAACATGTTTCAGTGT	2220
QY	2221	CCAAATTATTAGGTGGCTGACCAAGTTGGGTGAGTAACTGTTTGCANTATAGACTGGG	2280
Db	2221	CCAAATTATTAGGTGGCTGACCAAGTTGGGTGAGTAACTGTTTGCANTATAGACTGGG	2280
QY	2281	TGCTAAATTTCTGTGCTGTAAGCATTTTATAOCCATTTGTAAAGTCTTTAACTCTGGAA	2340
Db	2281	TGCTAAATTTCTGTGCTGTAAGCATTTTATAOCCATTTGTAAAGTCTTTAACTCTGGAA	2340
QY	2341	ACTTCGGGAAATTAATAAATGAATTCACAACTCTCAAAAAAAAAAAAAAAAAA	2396
Db	2341	ACTTCGGGAAATTAATAAATGAATTCACAACTCTCAAAAAAAAAAAAAAAAAA	2396

RESULT 2	AF429385	LOCUS	DEFINITION	ACCESION	VERSION
	AF429385	1785 bp	mRNA	linear	PLN 25-OCT-2001
			Hevea brasiliensis phosphomevalonate kinase mRNA, complete cds.		
	AF429385				
	AF429385.1	GI:16417947			

SOURCE	ORGANISM
Hevea brasiliensis.	Hevea brasiliensis.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eustersids I; Malpighiales; Euphorbiaceae; Hevea.	
1 (bases 1 to 1785)	

AUTHORS Hallahan, D.L. and Keiper-Hyuno, N.M.
 TITLE Genes involved in the biosynthesis of isopentenyl diphosphate in
 the rubber tree *Hevea brasiliensis*
 JOURNAL unpublished
 REFERENCE 2 (bases 1 to 1785)

AUTHORS	Keiper-Hrynko, N.M. and Hallahan, D.L.
TITLE	Direct Submission
JOURNAL	Submitted (05-OCT-2001) Central Research and Development, E. I. DuPont de Nemours Co., Wilmington, DE 19800, USA
FEATURES	Location/Qualifiers
SOURCE	1 1705

SOURCE
1:1:1705
/organism="Hewea brasiliensis"
/cultivar="RIM600"
/db_xref="taxon:3981"
/cell_type="latex"
121 1632
CDS

```

/codon_start=1
/product="phosphomevalonate kinase"
/protein_id="AAL18926.1"
/db_xref="GI:16417948"
/translation="MALVVASAPGKVLMTGGLTIERPAGIVLSTNNAFVIVKPIRY

```

EIKRDSNAAAMATDYKLTSPQIARRESLYKLSIKNLALOCVSSASRBNPVEDAOPAVAA
AAATLTLDKKNVNLNKLILLOGLDITLTGNDPYSYRNIEAGGLPIPESTALPSES
SITNPEDANQNCRPEYAKTGLSSAAMTAAVVAALPIHLGLVLDSSCKEKKFSIDI
DLVHIIAOTAHICIAQGRVSGSEFVSANYSGRHYRVESSAODAGKGIPLQEVET
SNILKGMDHERTMFSLPLMSLLLGEBGTGGSSTPSNVGLAKKWQKSDIYKSOETWFM

BASE COUNT	ORIGIN
554 a	351 c 406 g 474 t

KLSANSAALIFOTNITKLSAEHMDKACVTDSCSTKSEKIMIDATEAVEAVKAL
 LGTSNANMAIETNRMRQMEAGAVP IEDPSORLDTPTNMNGVILACVPGAGGCPDAV
 AVTIGDSGTNVAKAMSSLNVLALVREPNVLLSGDPRKREITTAFAVANI"

Query Match	35.1%	Score 841.2;	DB 8;	Length 1765;
Best Local Similarity	73.28;	Pred. No. 1.3e-214;		
Matches 1109; Conservative	0;	Mismatches 398;	Indels 9;	Gaps 2;

QY	684	AATGGCTGTGTGTGCTTCTGCTCTCGGAAAGTTTGTAGTACGTGAGGCTACCTTGTACT	743
Db	120	AATGGCTGTGTGTGCTTCTGCTCTCGGCTAAAGGTGTGATGATCGGGGTTCCTCATATT	179
QY	744	CGAAGAACCAATTCGAGGGCTGTGTGTAGTTCAAATGACGGTTTTACGGATTGTGAA	803
Db	180	GGAAGAACCCAAATGACGGGATTTGTACTCGACCAAAATGCTCATTTCTATGCAATTGGA	239
QY	804	GCCAAATCAGAAAGTGAAGCTGAAGCTTGGGCTATGGAATGGACAGATGTCAAATT	863
Db	240	GCCATATTACGATGAATCAAACTGTATGTTGGGATGGGATGGGATGATGTGAATT	299
QY	864	AACATCACCAAGCTCTCGAGGAAAGCATGTATAACTGTCTACTGAATCATATTGACTCT	923
Db	300	AACATCTCCCCAACTAGCAAGGAAAGCTTTGTACAAATTGTCTACGAAAAATTTTAGCTCT	359
QY	924	TCAATCTGTGTGCAAGTATTCAGAAACCCTTTGTAGAGCATGGCATACAGTATGC	983
Db	360	TCAATGTGTCTTTCAAAGTGCATCAAGGAACCACTTTGTGGAAACAGAGTGCATTTGC	419
QY	984	TATAGCTGCTGCTCATTTGGCAACGGAAG---ACAAAGATCATTTGCACAACTCTT	1040
Db	420	TGTAGCAGCTGCACATGCACAACTTGACCAAAATTAAGAAATGTCTTAAACAGCTACT	479
QY	1041	ATTGCAAGGCTTGGATATTAACATATTATAGCTGCACATGACTTTTACTCATATCGGAACA	1100
Db	480	CTTGCAAGGCTTGTGATATTTCATATTATATAGTACCAATGACTTCTATTATACCGAAATGA	539
QY	1101	GATAGAAATGGCTGGGCTTCATTTGACACACAGAAATCGCTGGTACCCTTGACCCCTTTCG	1160
Db	540	GATTTGAAGCATGTGGACTCCCTTTGACACCAAAATCATTTGGCTGCATCTCTCTTTTC	599
QY	1161	ATCAATTCATTTCAATAGCTGCGGAGTCAAAATGTGTATTTCCAGACCTGAAAGTGGCAA	1220
Db	600	CTCAATTCACCTTAAATGTAAAGGAAACAAATGGACAAATCTCAAGCTCGAGTATGCTTAA	659
QY	1221	AACGTGGCTTGTGCTTCTTCGACAGCAATGCAACAGCTGTGTGTGACGCTGTGTACATTA	1280
Db	660	AACGTGATGGTGTTCATTCAGACACACATGACCACTGCTGTATGTGTGTACTTCATCA	719
QY	1281	TCTTGGAGTGTGACCTTATCTGATCATCATGTAAAGAGAAAGTTTGGCTGTCTGATCT	1340
Db	720	CTTGTGATTTGGTGTGATCTTTCATCTCTCTTGAAGAAAGAAATT-----TTTCGATCT	773
QY	1341	AGATGTTATCATATGATGATGACAAACAGCTCATATGCTTTGGACAAAGGAAAGGTGTGAG	1400
Db	774	TGATTTGTGCAT	833
QY	1401	TGGGTTTGTATGTCAGCTGTGCTGTCTATGTGAAGTACAGCGTTATGTTGCTCTCTCAGA	1460
Db	834	TGGATTTGATGTAGTTCTGTCAGTATTAATGAGCTCATCATGATCGTCCCTTCTCTCAGA	893
QY	1461	AGTCTTGTATTCGTGCAGGTTTCCAGTACAGTACAGTCTGCATTAATTAAGATTAATTTTAT	1520
Db	894	AGTCTTGTCTCTGCTCAGGATCTGTGGAAGGAATTCATTACAGAAATGTCATTTCTTAA	953
QY	1521	AATTTTGAAGGAAATGGACAAATTAAGAAACTGAGTTCTCTTTTACCACACATGTATGAA	1580
Db	954	CATCTCTTAAAGAAATGGGACCATGAGAGCATATGTTTCTCTTCCACCATTTGATGAG	1013
QY	1581	TCTTTTCTTTTGGAAACCTGGAAGTGTGTGATCTTCACACATCATATGTATAGTGCAGT	1640
Db	1014	CTGTGTACTTAAAGTACACAGGAACTGTGAGATCTTTCACAGCCATCAATGTATAGTGTCT	1073
QY	1641	AAAGAAATGCAATGTGTGATTCACAGAAAGCAGACAGAAATCTGCAGAAATTTTGTGCA	1700
Db	1074	AAAGAAATGCAATGTGTGATTCACAGAAATCTGTGATTCACAGAAATCTGTGATTCACAG	1133
QY	1701	TGCAAAATTTTGAACCTGGAACCTTAAAGTCTGAGCAAAATTTAGCTTAAAGACCACTG	1760


```

|||||
Db 28358 AATGTTGGTGCAGTAAAGATGGCAAAATGCTGATCCAGAGAGACAGAGAAAACGTG 28299
QY 1686 GCAGAAATTTGTCAGATGCAATTTAGAACTGGAACCTAGCTAAACGATCGAGCAAT 1745
Db 28298 GCAGAAATTTGTCAGATGCAATTTAGAACTGGAACCTAGCTAAACGATCGAGCAAT 28239
QY 1746 AGCTAAAGACCACTGGGATGTTATCTAGAGATCATTAAGTCTTGAGTGCTTACTTC 1805
Db 28238 AGCTAAAGACCACTGGGATGTTATCTAGAGATCATTAAGTCTTGAGTGCTTACTTC 28179
QY 1806 TGAAA----- 1810
Db 28178 TGAAGAGATATTTCTCTTCTGAGAGATTTGATTTGTAAGATATTTTACATGAAGGG 28119
QY 1811 -----AGTGGTGTATCATGCTACTGAAC 1835
Db 28118 CAGTTTGTTCTATTTGACTTTTGTATCTCATGTTTACTGAGTGGTGTATGCTACTGAAC 28059
QY 1836 AATCAACGAAGCCATTTATTAAGAACTTTAGAGCAAGAGAACCTATGTTGAGATCAG 1895
Db 28058 AATCAACGAAGCCATTTATTAAGAACTTTAGAGCAAGAGAACCTATGTTGAGATCAG 27999
QY 1896 AATCTATGCTCAGATGGGTGAGGGGCTAGCGTCC----- 1934
Db 27998 AATCTATGCTCAGATGGGTGAGGGGCTAGCGTCCGTTCCGTTAGATTTTGTCTGT 27939
QY 1935 ----- 1934
Db 27938 TTGATTTGCTTCAATCTCAACGAGCAATTTTGATGATATTTACTTATGCTGTTCAT 27879
QY 1935 ----- 1934
Db 27878 AATAATCTGCTGCTGCTGATAGATCTTAACAGCATATGATATPACGGGTTCCACAAT 27819
QY 1935 -----G 1935
Db 27818 TCCGCAATTTAGAACAAATTTGACTTCATCATTAATTTGACCTGAGGAGCAATGGAATTCAG 27759
QY 1936 ATAGAGCCTGATCTCAAACTCAAACTTTTGATTTCTACAATGATGCTGAAGAGTTCTA 1995
Db 27758 ATAGAGCCTGATCTCAAACTCAAACTTTTGATTTCTACAATGATGCTGAAGAGTTCTA 27699
QY 1996 CTTCGCTGCTGCTGAGCTGCTGATTTGATGCTCAATTTGCAATCTTAAAGGGAT 2055
Db 27698 CTTCGCTGCTGCTGAGCTGCTGATTTGATGCTCAATTTGCAATCTTAAAGGGAT 27639
QY 2056 TCCGGACCAAACTGACCCAGGAGTGGATTCGACAAATGTTTGGCTTGTGTGAGA 2115
Db 27638 TCCGGACCAAACTGACCCAGGAGTGGATTCGACAAATGTTTGGCTTGTGTGAGA 27579
QY 2116 GAAGATCCACATGCGCTTTCCTGAGAAAGTGTGATCCAGAACACATGATTAATCTCA 2175
Db 27578 GAAGATCCACATGCGCTTTCCTGAGAAAGTGTGATCCAGAACACATGATTAATCTCA 27519
QY 2176 GCGCTTTCATCAATTCACCTTGAAGTAAACAATTTGTTCACTGCTCAATTTAGGTGC 2235
Db 27518 GCGCTTTCATCAATTCACCTTGAAGTAAACAATTTGTTCACTGCTCAATTTAGGTGC 27459
QY 2236 GTACCAAGTTCGGTGAATATCTGTTTGCATATAGACTTGGGTCTAAATTTCTTGG 2255
Db 27458 GTACCAAGTTCGGTGAATATCTGTTTGCATATAGACTTGGGTCTAAATTTCTTGG 27399
QY 2296 TGTAGCAATTTTATACCAATTTGAAGTCTTTAACTCTTGAAAACTTCGCGGAATA 2355
Db 27398 TGTAGCAATTTTATACCAATTTGAAGTCTTTAACTCTTGAAAACTTCGCGGAATA 27339
QY 2356 AAATAAGTGAATTTCAATCTTCTCA 2382
Db 27338 AAATAAGTGAATTTCAATCTTCTCA 27312

```

```

AX41245
LOCUS AX41245. 728 bp DNA linear PAT 02-JUL-2002
DEFINITION Sequence 4 from Patent EP1209236.
ACCESSION AX41245
VERSION AX41245.1 GI:21690241
KEYWORDS
SOURCE
ORGANISM
Gossypium hirsutum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
Rosidae; eustosids II; Malvales; Malvaceae; Gossypium.
upland cotton.
REFERENCE
1
AUTHORS Weissen, R. and Lechelt-Kunze, C.
TITLE Phosphomevalonate kinases from plants
JOURNAL Patent: EP 1209236-A 4 29-MAY-2002;
BAYER AG (DE)
FEATURES
location/Qualifiers
source 1..728
/organism="Gossypium hirsutum"
/db_xref="taxon:3635"
BASE COUNT 213 a 151 c 144 g 219 t 1 others
ORIGIN
Query Match 17.8%; Score 425.4; DB 6; Length 728;
Best Local Similarity 76.1%; Pred. No. 6; 3e-103;
Matches 554; Conservative 0; Mismatches 162; Indels 12; Gaps 2;
QY 784 CCGTTTACCGGATTTGAGACCAATCAACGAAGAACTCAAGCCTGAAAGTTGGGATG 843
Db 1 CGTTTACCGCTATTTGTAAGCCATATCATGAAAGCTATCAAGCCTGAAAGTTGGGATG 60
QY 844 AATGACAGATGTCAAATTTAATCATCACACAGCTCTGAGAGAAAGCATGATTAAGT 903
Db 61 TCTTGGACCCATGTCACAGTCAACATCTCTCAGCTTCCAGAGAAAGCATGATTAAGT 120
QY 904 TCCCTGATCATATTTGATCTCTGATCTGCTGCTGCAAGTATCAAGAAACCCCTTGTGA 963
Db 121 TCTCGGAACATTTTAACATCTGATGCTGATCTTCAAGTATCAAGAAACCCCTTGTGA 180
QY 964 GAGCATCGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
Db 181 GAAATGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
QY 1021 GAATCATGCAAACTCTTATTCGAAGCTTGTGATATTAACAATATTAGCTCCATGAC 1080
Db 241 GAGCATTTAGATTAACATCTTCAACAGCTTGTGATATTCAGATCTTGAAGCTTGA 300
QY 1081 TTTTACTCATATGCAACCGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
Db 301 TTTTACTCATATGCAACCGATGATGATGATGATGATGATGATGATGATGATGATGAT 360
QY 1141 GGTACCTTTCACAGGTTTGTGATCAATCAATTCATGCTCGAGAGTCAAAATGGTGAAT 1200
Db 361 GGTACCTTTCACAGGTTTGTGATCAATCAATTCATGCTCGAGAGTCAAAATGGTGAAT 420
QY 1201 TCCAAAGCTGAAGTAGCAAAATGCTGTTAGTCTTTCGACGAATGACAAACGCTGTG 1260
Db 421 TCCAAAGCTGAAGTAGCAAAATGCTGTTAGTCTTTCGACGAATGACAAACGCTGTG 480
QY 1261 GTTCAGCTCTGTTACTATTTCTTGGAGTGTGACCTATCTGATCCATGATGAAG----- 1315
Db 481 GTTCAGCTCTGTTACTATTTCTTGGAGTGTGACCTATCTGATCCATGATGAAG----- 1315
QY 1316 -----AAGGAAGTTTGGCTTTCGATGATGATGATGATGATGATGATGATGATGATGAT 1371
Db 541 CACCAAGAAATTAAGATTTCCACAGATCTGATATTTGATATGATAGCTCAAAAGTGC 600
QY 1372 CATTTCTTCACAGGAAGGATGGAAGTGGTTGATGATGATGATGATGATGATGATGATGAT 1431
Db 601 CACTGATTTCCCAAGTAAGTTGGCAGTGGCTTTGATGATGATGATGATGATGATGATGATGAT 660
QY 1432 AGTCAGCTTATGTTGCTTCTCTCCAGAAAGTCTTGATATTTGCTCAAGTTGATGATGAT 1491

```

RESULT 4

Db 661 AGTCAGCGTATATGCTCTTTTTCACCAAAAGCTTCTGCTCAGGCTGCANTGAAA 720
 QY 1492 GGTCGCC 1499
 Db 721 GGGATGCC 728

RESULT 5
 AX441244
 LOCUS AX441244 611 bp DNA linear PAT 02-JUL-2002
 DEFINITION Sequence 3 from Patent EP1209236.
 ACCESSION AX441244
 VERSION AX441244.1 GI:21690240
 KEYWORDS
 SOURCE barrel medic.
 ORGANISM Medicago truncatula
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
 Medicago.

REFERENCE
 1
 Meisner, R. and Lechelt-Kunze, C.
 Phosphomevalonate kinases from plants
 Patent: EP 1209236-A 3 29-MAY-2002.
 BAYER AG (DE)
 location/Qualifiers

FEATURES
 1. 611
 /organism="Medicago truncatula"
 /db_xref="taxon:3880"

BASE COUNT 167 a 135 c 121 g 188 t
 ORIGIN

Query Match 12.5%; Score 299.2; DB 6; Length 611;
 Best Local Similarity 71.7%; Pred. No. 5e-69;
 Matches 407; Conservative 0; Mismatches 158; Indels 3; Gaps 1;

QY 691 GTTGTGCTCTCTGCTCTGGGAAAGTTTGTGATGACGAGGCTACCTGTACGAGAG 750
 Db 44 GTGGTCTCTCTGCTCTGGGAAAGTTTGTGATGACGAGGCTACCTGTACGAGAG 103
 QY 751 CCAATGACGAGGCTTGTGTGATGACGAGGCTTGTACGAGGCTTGTACGAGGCTTGTACGAGGCT 810
 Db 104 CCAATGACGAGGCTTGTGTGATGACGAGGCTTGTACGAGGCTTGTACGAGGCTTGTACGAGGCT 163
 QY 811 AACCAAGACGAGGCTTGTGTGATGACGAGGCTTGTACGAGGCTTGTACGAGGCTTGTACGAGGCT 870
 Db 164 TATCTGCAAACTAACTGATCTTGGGCTTGGCTTGGCTGATGATGATGATGATGATGATGATGATGAT 223
 QY 871 CCACAGCTCTGAGAGAAAGCATGTATTAATCTGATGATGATGATGATGATGATGATGATGATGATGAT 930
 Db 224 CCACTGATCTGAGAGAAAGCATGTATTAATCTGATGATGATGATGATGATGATGATGATGATGATGAT 283
 QY 931 GTGCTCAGTGTATCAAGAAAGCATGTATTAATCTGATGATGATGATGATGATGATGATGATGATGATGAT 990
 Db 284 GTTGTGCAAACTAACTGATCTTGGGCTTGGCTTGGCTGATGATGATGATGATGATGATGATGATGATGAT 343
 QY 991 GTTGTGCAAACTAACTGATCTTGGGCTTGGCTTGGCTGATGATGATGATGATGATGATGATGATGATGATGAT 1047
 Db 344 GCGGCGATGAGAGAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 403
 QY 1048 GGTCTGATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1107
 Db 404 GGTCTGATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 463
 QY 1108 TCGGCT 1167
 Db 464 AGACAGCAGCTCCCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 523
 QY 1168 ACATTCATGCT 1227
 Db 524 TCTTCAATATGAT 583
 QY 1228 TTAGGCTTCTGAGCAGCATGACACAG 1255

Db 584 TTGGGCTCATCTGCAGCATGACCAACCG 611
 RESULT 6
 AB011480
 LOCUS AB011480 40548 bp DNA linear PLN 27-DEC-2000
 DEFINITION Arabidopsis thaliana genomic DNA, chromosome 5, p1 clone:MP17.
 ACCESSION AB011480
 VERSION AB011480.1 GI:2924730
 KEYWORDS
 SOURCE Arabidopsis thaliana (strain:Columbia) DNA, clone_1lb:Mitsui P1
 Clone:MP17.
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 1 (sites)
 Kaneko, T., Kotani, H., Nakamura, Y., Sato, S., Asamizu, E., Miyajima, N.
 and Tabata, S.
 Structural analysis of Arabidopsis thaliana chromosome 5. V.
 Sequence features of the regions of 1,381,565 bp covered by twenty
 one physically assigned P1 and TAC clones
 DNA Res. 5 (2), 131-145 (1998)
 98344145
 2 (bases 1 to 40548)
 Nakamura, Y.
 Direct Submission
 Submitted (02-MAR-1998) Yasukazu Nakamura, Kazusa DNA Research
 Institute, Department of Plant Gene Research, 1532-3, Yana,
 Kisarazu, Chiba 292-0812, Japan (E-mail: ynakamu@kazusa.or.jp,
 Tel:81-438-52-3935, Fax:81-438-52-3934)
 Address for correspondence: kaos@kazusa.or.jp
 For the latest information on annotation of this clone, please see
 http://www.kazusa.or.jp/kaos/cgi-bin/seq_graph.cgi?c=mp17
 Genes with similarity to proteins in the databases are described in
 'product' or 'note' qualifiers. Genes that have no significant
 protein similarity are described as 'unknown protein'.
 The software programs used to predict genes include: Graal
 (Informatics Group, Oak Ridge National Laboratory,
 http://compbio.ornl.gov/Graal-1.3/),
 GENSCAN (Chris Burge, MIT, http://CCR-081.mit.edu/GENSCAN.html),
 NetGene2 (S.M. Hebsgaard, et al., CBS, Technical University of
 Denmark, http://www.cbs.dtu.dk/services/NetGene2/) and
 SplicePredictor (Volker Brendel, Stanford University,
 http://greml.inl.zoology.utoronto.edu/cgi-bin/sp.cgi).
 Genes encoding tRNAs are predicted by tRNAscan-se
 (Sean Eddy, Washington University School of Medicine, St. Louis,
 http://genome.wustl.edu/eddy/tRNAscan-se/).
 This sequence may not be the entire insert of this clone. It may be
 shorter because we remove overlaps between neighboring submissions.
 The 5' clone is MW3 and the 3' clone is MCM23.
 location/Qualifiers

FEATURES
 source
 1. 40548
 /organism="Arabidopsis thaliana"
 /strain="Columbia"
 /db_xref="taxon:3702"
 /chromosome="5"
 /clone="MP17"
 /clone_1lb="Mitsui P1"
 /complement(3187..4899)
 /note="contains similarity to unknown protein
 gb|AA025785.1
 gene_id:MP17.2"
 /codon_start=1
 /evidence=not_experimental
 /protein_id="BAB11219.1"
 /db_xref="GI:1017788"
 /translation="MASLFSRLSSQSLINIFLFLFPAQSQTPPPSGIQTL
 NSPAGSDSDSGSLASLDHRSKSYRSKSGPGGYIDYKIFCIRGSGVLGH
 LYSAMFLFVFLIGDRAAFYFCSLDSLVKLSPTMAGVLTLSLNGAPDLFSSV
 VSPFRSNKDGFLNLSIGAFVSPVCGTCTGSRDAIDNSIRIVPVLTALV
 CCLGLIFIGKVTIWAALCTIYLIVGLFSLVSHFDRKKRMSDQILRSREDLAEKG

CDS
 8366..8686
 /note="contains similarity to plastid-specific ribosomal protein 6"
 /gene_id="MP17.3"
 /codon_start=1
 /evidence=not experimental
 /protein_id="BAB11220.1"
 /db_xref="GI:1017788"
 /translation="MSVSAIFGTVTAASPVLRQFVPLKLGNGGLGMVTECSSRP OKSTVAHNRKTPKKTQPMWDIRKRPVAPLPPLPAEMSPFTLASDDGAATAAGDLY SGAA"
 complement(join(9595..10290,10376..11266,11359..11673,11914..13096,13164..13645))
 /note="gene_id:MP17.4"
 /evidence=not experimental
 /product="disease resistance protein"
 /protein_id="BAB11221.1"
 /db_xref="GI:10177889"
 /translation="MTSSSWVKTGDETPQDQVFNFRGVELRKNFVSHLEKLRKG INAFIDTEMGOELSVLERIEGRKIALAIFSPRYTESKWLKELAKKERTEOEL VYPIFYQVPTVKEIKGDFGDKRELDKTKKEMKALQVPLPTGIVDEL SVSDEDEVINIIIRKVEIIRNRSEGPSPKCALPDRHOKRHEFPIELRIKOLE EKLRGSDETRTIGVGMGKTTTALMLYKEMVRELRYLIRDIJHEASEEGRN YLATRFLOGLAKVENANISVQAHAHAYKQDLLETKVYLIDNVSKDDVALLGERN WIKRGSKITITTSKSLMTQSLVNDYEVPLSDKALNFIIRYAFDGEAGALPGGQ GNPRLSKDFVHYTKGNPLALQMLGKELGKDESHWGLNALMDHNSPPOSICIM LORWEGSYKALSOKEKDALDIAFCRSODENVASLSDSDPSNLELVKFMKINI YAGKDMHDTLLSKELGREATDRKGRHRLMHHTTIIAVLDKRRKGSNTISFID LSDIRKMCFTYRHAFAAMRDLRYIKRYTSRPOCESDILKINPEGLIPLNEVRIH WIKPFLKVPDPNPNGLVLDKLPYSEIRVWEDNDARLKWNLNHSKNTLTLG GAQNLQDELNLEGCTALKEHVMEMKFLVELNKGCTSLKSLPEIQLISLKTLLS GSKRTFQVVIDKLEALYLDGTALKEIPCDIGRLQRLVLMKGGCKLRKLRPDSIG LKALEELISGSKLINEPFTGWNMSRELILLDETAIKDMKILSVRLCLNKIKI SRPLDLNKSOLQMLHLYCKNLTHVPLPOLPNLOYLVNMGSSLSKTAVLKPISPM KHVNSFTITNNELEOAKREIIVYAEKCHILASAKRDESCYELTCTSPGC EMPNSFSDHATISWVEFELPRMNNRSLGIALCYVYSKNCVSHANLIVKSCEDN GEGSSSITWAKVSLIEDNOEIVESSHVFIGYINCDFIKLVGGGGPKCAPTKAS LEFSVATGGGATLEVLKSGSFVEFEPENRPPSRNDVAGKVIINTPSANCGFK DOAKNESPKGMQVYIEMNSTINISEAHSQKTGFGSTLLC"
 join(117564..18018,18326..19451,19542..19838,19924..20694,20792..21898)
 /note="gene_id:MP17.5"
 /codon_start=1
 /evidence=not experimental
 /product="disease resistance protein-like"
 /protein_id="BAB11222.1"
 /db_xref="GI:10177890"
 /translation="MEPPARVTPSIKADCSHSVNICETVLSHLSLAAREG ISVPDAGCLOETKFPSTKQNPPLTDGARVLYVSDVEFEDDPMPFKLKIYQON NCHVAVPEYGVDSLTRYGMANSWLEAKLTSHOSKFVYIILSFVKPILSNVLT DSELVEIVRDVYGLYPAERGVIAARLEIKLYKORHDIRSIGIWMKPIKGTLL AKAVPNHMSSTDASCFTIENPEAFHKEGLHLEKRIKILKDEFDIESSYIMRTL HRDKLYDKRIILVLDVDRSLLAESFLKRLDMFGSSLLITISVDKOVARPOLINOIY TVOGLNVHEALQFSOSGVINPEQNDKLSMKYLDVYDNGPMLAISITVGRILMKS EMETAFELKRCPPRIKIDVLYKNAYSALSDNKENYIADIAFFKGETVYVQGLEES HYFPRILADVLDKCVLTISNTVQNNNLIQDTCOEFGELTCTRMERPIRIRLL EYDELGSETKAMPKSGVAHIESIFLDTSNVKEVDHAFKNNFLKPIKATNSC SKYISGLNPKGLDLPYELRLHWHENYFLOSLPODFGHLVKLSMPSQLAKITNC YDLVLYKRLILSHSLOVECDILYAOINIELIDOGCTGQRPTDSQLOMLRYVNL SGTETKCSGYVPNIIEELHOGTRIRELPIFNATHPPVKVLDKRLKMLNLFNSVE HIDECVYNLATVTSNNHVMGLVCLNMKTCNLKQLPWVSESLKAVIYLSGCSLE KIMGFPRNLKRLYVGATIRELPQPLNSLEFINAGCKLKLINLDFEDLPHFIFSN KYRFSQVYIAEVKGLVASLARAKOEELIKAPVEITCIPMTRKTSRLOAGRAN TDLVPMQKPIISGFSVYVFDODYHNHVGRLIKVWKTGMNNOPIADIVRFDCCW APTEAPKVYADHIFVLYDTKMHPSDESEHHSIMAHVEGVKFEETHTVSGENLPGASCKV TECGVEVITAACTGTSVSGITIESETITTIIEKDTIIDEEDTPLLSRKEETENRNS

CDS
 22878..22949,23030..23148,23223..23281,23363..23709)
 /note="emb|CAB16816.1|
 /gene_id="MP17.6"
 /similar to unknown protein"
 /codon_start=1
 /evidence=not experimental
 /protein_id="BAB11223.1"
 /db_xref="GI:10177891"
 /translation="MNEISDLGMHNRPCYKELRHPNVCYCEKKIPRTAEGLYNHP FMWEYCPSHDDGTPKCCSCERLEGTQYVWLDVFNLCRCDMSALMDSDECPL HFEIRREFEGHLMKIEEPVYLVENKALNKAKEKIDKODGCLMAYRGICLSEQ IYTSQGVRRMLNKOIIDTVTESQVSKCVTAIILIGYPRLLTYGILAHENMHA YLRNLYRMLNVNLEGLCOVGYMVECOYVFPATYASSSSRTPLSTTSKV DPSDEPKRLVNCCKHOIETDESFPDGGFRKYNKMAANNHSLKOTLKILISIKPQ YSKL"
 join(25693..26888,27128..27239)
 /note="contains similarity to microfilin-like-associated protein 1"
 /gene_id="MP17.7"
 /codon_start=1
 /evidence=not experimental
 /protein_id="BAB11224.1"
 /db_xref="GI:10177892"
 /translation="MSVTVAGVSAVAAREKLGIGIGQTVRRVYMPKAPMAEAE DDVVRQKQVSLDRAPKNDQGVAKADDPRLAKTVKREVRADRAHRRIOAEI IETFEESNEDSDDEDDDEDAIRERIRKNIIRAOEALPLLEDDEIOEE EEESESTETDSEDMDGIMIKIPYVYKAEVOTIARENLAEELAEELARKLEOR KLETOQIVEEVKEDEIRKNILLEANIGDVEOTIDARENLAEELAEELARKLEOR RDAREMLAREEIEKRLMTDQERDMWRKKNPKLSAOPKRWMPKRYKRGAFQ ADPDEAGSAGEDVIOEDFSAPTGEDRLDKSLIPVYGVKHFGRSGRTKWHLVNED TDWMSWPTSNPLREKYNKKAAGAPAKPKSKKMKWES"
 join(28455..31403,31762..32841)
 /note="gi|AC35233.1|
 /gene_id="MP17.8"
 /similar to unknown protein"
 /codon_start=1
 /evidence=not experimental
 /protein_id="BAB11225.1"
 /db_xref="GI:10177893"
 /translation="MTSDSERVQIRLPMITRTYKICNHPPLGFAFLYTLH RAYTLFLAPLVASPVLYCTFVLTGILSFSGPNIPLEIKDEIPIEAPLPTESRD ANCVTVDRGDSFTVESVGAKEVLEEDGNDABRLVDSQSEVEDDGRPDYRLV DETLDEIKDTHVREFEKAFIILDEVEKGRDEKILENGDTAEOSRTGSLYERMD OMDVSPSPMRMRHEEDDDADRDSDSDGAESSSPASMTDIIPLMDELPL ILSEAPTRCTVDCGSDASBGRPHSSSPGEMSPDSRSHGDEEDNEDEDEE EDEBEKKEKEDDESKSATIKWTEADQNRVMDLSLELRNORLENLARRARRNM RLMAERNLIDFDSADIPEFMPPISTARHNPVSYDSYDMPIZGSAPSIARRNPF DLPEPNEKRPDLKGDGFQEEFSQPKPMPRRHSFVSGSMGLGPHRDLREPV LRLANEGESYVPERFOLSEVESKYSIPDTESVCTVLEDEKEDENNADRETKIA KYDWSNDNEENNHSASDHENSHASDDEKSHSSSDSPEDVADDSKLLHHVAAE IYVSGETHEDSDMMGEFTSDKGLDEVSDDSKISSEKEKIRDISDEAMLSVOY VDLHEIAGASSLPRELEITNARGVEDYIHNDKAREESIITAHPSIDESALHYLGC LDDGHEEPVYSSPPSGSRFSPSSVSDVAPDLPEKKEGELEENVEEKERYISEI GVEEHTSENETERTSSEGENSMHVTGASLVMRHSPLPEESDCLVADIAETSNK SVVEELIYEEEAOKOKDEVSPTFADIPDISVSLSGAAYEVETHFNFNEDVAOL EDEPVSLHVADEETHNDQMDIIEVDVSNASAVNWSSETSPESDERELTVSDSVV

Query Match 9.5%; Score 227; DB 8; Length 40548;
 Best Local Similarity 95.9%; Pred. No. 1,1e-49;
 Matches 233; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 18 GGGCGACCTTCTCTTCTTAAAGCAACACATATATAGAACAACTGGGAGAGA 77
 |||
 Db 24896 GGGCGACCTTCTTCTTCTTAAAGCAATATATATAGAACAACTGGGAGAGA 24955
 |||

QY 78 TGAAGATGAGATGGTGAAGAACAACCGTATACCGTTCTGAGAGGTGGCGAACCG 137
 |||
 Db 24956 TGAAGATGAGATGGTGAAGAACAACCGAATACCGTTCTGAGAGAGGTGGCGAACCG 25015
 |||

OY 138 AACGACCCGTAACCGAAATCTCAAAAGAAATTCGCCATCGCTTGTCTACTGTTCAAA 197
 DB 25016 AACGACCCGTAACCGAAATCTTGAAGAATTCGCCATCGCTTGTCTACTGTTCAAA 25075
 OY 198 ACCTCGGTGCGGACCGAAACGCTGTTTTCGGTGGGTTTCTCGGTTTCTTCC 257
 DB 25076 ACCTCGGTGCGGAAACCGAAACGCTGTTTTCGGTGGGTTTCTCGGTTTATTC 25135
 OY 258 GAA 260
 DB 25136 GAA 25138
 RESULT 7
 LOCUS AX441246 571 bp DNA linear PAT 02-JUL-2002
 DEFINITION Sequence 5 from Patent EP1209236.
 ACCESSION AX441246
 VERSION AX441246.1 GI:21690242
 KEYWORDS
 SOURCE Monterey Pine.
 ORGANISM Pinus radiata
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Coniferales; Pinaceae; Pinus; Pinus.
 REFERENCE 1
 AUTHORS Meisner, R. and Lechelt-Kunze, C.
 TITLE Phosphomevalonate kinases from plants
 JOURNAL Patent: EP 1209236-A 5 29-MAY-2002;
 BAYER AG (DE)
 FEATURES
 source 1..571
 /organism="Pinus radiata"
 /db_xref="taxon:3347"
 BASE COUNT 161 a 129 c 136 g 145 t
 ORIGIN
 Query Match 5.6%; Score 135.2; DB 6; Length 571;
 Best Local Similarity 66.5%; Pred. No. 5.7e-25;
 Matches 210; Conservative 0; Mismatches 103; Indels 3; Gaps 1;
 OY 683 AATGGCTGTGTGCTTCTGCTCTCGGAAAGTTTGTATGCTGAGGCTACCTTTGAC 742
 DB 242 ACATGGCTGTGTGCTTCTGCTCTCGGAAAGTTTGTATGCTGAGGCTACCTTTGAC 301
 OY 743 TCGAGAGGCAATGAGGCTTGTGTGATCAAAATGCAAGGTTTACGGGATTTGCA 802
 DB 302 TTGAGAGGCAATGAGGCTTGTGTGATCAAAATGCAAGGTTTACGGGATTTGCA 361
 OY 803 AGCCAATCAAGCAAGAGTCAAGCTGAAAGTTGGGATGAAATGACAGATGTCAAAT 862
 DB 362 AGCCAATGAGCAATGAGGCTTGTGTGATCAAAATGCAAGGTTTACGGGATTTGCA 421
 OY 863 TACATCACCACAGCTCTGAGAGAAAGCATATTAATGCTGACTGAATCAATTGACTC 922
 DB 422 TACATCACCACAGCTCTGAGAGAAAGCATATTAATGCTGACTGAATCAATTGACTC 481
 OY 923 TTCACATGCTGCTGCTGCAAGTCAAGAAACCCCTTTAGAGGATGATGATC 979
 DB 482 TCCAAATGCTGCTGCTGCAAGTCAAGAAACCCCTTTAGAGGATGATGATC 541
 OY 980 ATGCTATAGCTGCTGC 995
 DB 542 TTGCTGTTGACAGCTGC 557
 RESULT 8
 LOCUS AX461344 2004 bp DNA linear PAT 08-JUL-2002
 DEFINITION Sequence 273 from Patent WO0198480.
 ACCESSION AX461344
 VERSION AX461344.1 GI:21726552
 KEYWORDS
 SOURCE thale cress.

ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 REFERENCE 1
 AUTHORS Budworth, P., Brown, D., Chang, H.S., Zhu, T., Han, B., Wang, X. and
 Cooper, B.
 TITLE Promoters for regulation of plant gene expression
 JOURNAL Patent: WO 0198480-A 273 27-DEC-2001;
 Syngenta Participations AG (CH)
 FEATURES
 source 1..2004
 /organism="Arabidopsis thaliana"
 /db_xref="taxon:3702"
 BASE COUNT 751 a 341 c 264 g 648 t
 ORIGIN
 Query Match 5.1%; Score 123.2; DB 6; Length 2004;
 Best Local Similarity 71.2%; Pred. No. 9.3e-22;
 Matches 195; Conservative 0; Mismatches 68; Indels 11; Gaps 2;
 OY 271 TAGTTGGTTTATTTTCACAGTTTGTCTCTTTTCATCGGCGACAGAGCTGCA 330
 DB 512 TAGTTGGTTTGGGATTCACAGTTTGTCTCTCTCTGACGGCGAC-TTTTACGGC 454
 OY 331 GTTCTGTCAAAAGCTTAACAGTCCGACTCGAGCTCGACATAAGACAGACAGCG 390
 DB 453 GATTTATCTCTCCATATCATGCTGTCTGATCATTAAGACAGACAGCGCA 394
 OY 391 ATTGTGTG-----TAGATGACGCGGAGAGCTGTGATCCGTCTGATGACGG 440
 DB 392 ATTGTGTGCGATCCGCTCGATCAACGCGGATGATGATCCGTCTGATGACGG 334
 OY 441 AAGATACGTTGATCCGCTTTCGATCCAAATGCGAGATTTGAGATCAATGAGAA 500
 DB 333 CGATAGTGTGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTC 274
 OY 501 TTGCATTAATCAATCTCCAAATCTCTCTGAG 534
 DB 273 ATTATGATCAATCTCCAAATCTCTCTGAG 240
 RESULT 9
 LOCUS AC004705 89840 bp DNA linear PLN 11-MAR-2002
 DEFINITION Arabidopsis thaliana chromosome 2 clone F26C24 map m3398, complete
 sequence.
 ACCESSION AC004705
 VERSION AC004705.3 GI:20197260
 KEYWORDS HTG.
 SOURCE Arabidopsis thaliana.
 ORGANISM Arabidopsis thaliana.
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 REFERENCE 1
 AUTHORS Rounsley, S.D., Kaul, S., Lin, X., Ketchum, K.A., Crosby, M.L.,
 Brandon, R.C., Sykes, S.M., Mason, T.M., Kerlavage, A.R., Adams, M.D.,
 Somerville, C.R. and Venter, J.C.
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 89840)
 TITLE Direct Submission
 AUTHORS Town, C.D. and Kaul, S.
 JOURNAL Medical Center Dr., Rockville, MD 20850, USA
 REFERENCE 3 (bases 1 to 89840)
 TITLE Direct Submission
 JOURNAL Submitted (27-FEB-2002) The Institute for Genomic Research, 9712
 Medical Center Dr., Rockville, MD 20850, USA, cdtown@tigr.org
 COMMENT On Apr 18, 2002 this sequence version replaced gi:6598440.
 FEATURES
 source 1..89840
 Location/Qualifiers


```

repeat_region      10403..10423
                    /rpt_family="AT_rich"
gene               10587..18487
                    /gene="At2g14720"
                    /note="T6B13.2: identical to GB:079960: contains a calcium
binding EGF like domain signature; supported by cDNA:
g1_1737219_gp_079960_1.At079960"
mRNA               12199..12282,12422..12498,12566..12647,12995..13077,
                    13170..13236,13328..13456,13617..13726,13904..14249,
                    18472..18487)
                    /gene="At2g14720"
                    /join(10724..11027,11107..11799,11889..11977,12056..12125,
                    12199..12282,12422..12498,12566..12647,12995..13077,
                    13170..13236,13328..13456,13617..13726,13904..14002)
                    /gene="At2g14720"
                    /codon_start=1
                    /product="putative vacuolar sorting receptor"
                    /protein_id="AAM15052.1"
                    /db_xref="GI:20197382"
                    /translation="MKQLCYLPMLLLSLVSPFNEARFVERKNSLSVTSPIKNGT
HDSALNFGIPQYGSAGTAVVYKPEKQSKSEFSDPSISKQSGALPFLVDRD
CFPALKVMNOKAGASAVLVADNDEPLTMDTPEEDVSATKIENTITPISALVTGF
GELKKRAISGDMVNLMDPRAVPPHDDPVEYELMTNSDEGVKCDMLMEYKQEK
GAQILIEKGFOTFPRPHYITWCPHATLSRQCKSCINCRICAPDEDESSGIDG
KDVVENLRQLCVYKVANETGKFWVMDVTDQIRKPKERKYNMDCASVILSGI
DSRRIDKCMQDPADLDNPLKEBDAQVGRGTGDTILPTLVNRRQYRGKLENSA
VLKALCSGEESTEPALICLSTDMETNECLDNGGCMADANITACDKRGKVCVP
IYGVREKGGDYSCHESGPRCTINNGCGMHEBRDHSACVADKQSVCECPGPK
GDGKRCEDI NECKEKACCCPCSCKNMGSYSCGDLIMRPHDCTISKTSQV
KSMVAWVLMISLGLAAGATVYKTRLMQYMDSEIRAIQAOYMPLDOSPEVPHNIN
DERA"
gene               complement(14856..15944)
                    /gene="At2g14710"
                    /note="T6B13.1: predicted by genscan and genefinder"
                    /gene="At2g14710"
                    /complement(14856..15944)
                    /gene="At2g14710"
                    /codon_start=1
                    /product="hypothetical protein"
                    /protein_id="AAM15050.1"
                    /db_xref="GI:20197380"
                    /translation="MAHLKMLPWEILIEILSRVPPKSLVFRVSKOMALFEDKFTI
NNHKTFRFLATKRSKIYSVSDPVIYVRELPGIPELESLENNLDCWELLVCYKN
NGAVVNPWLGOSRMKIOPSLNHSFVDFGVYNNKKYKMAFGLMKIYDFSSDVKN
HKSSTNSNTNINVOTAVCLNGLNLYVCCRERKIDPLCYIHKFGSNDIYVRFENLP
FGNHGDALVLCVFRGDRSLKOCWTKIEFVTEKINHNHNGSDIYVKNMFTS
SPMLPLVETVSRKSNPSYFTEGRKRVYCSDNNGHMIYILGSKLSKRIECVVD
WPLHCTFVSLVYVPAPCRREQAELOY"
repeat_region      complement(16330..16365)
repeat_region      /rpt_family="AT_rich"
repeat_region      16650..16767
                    /rpt_family="AT_rich"
repeat_region      16920..17003
                    /rpt_family="AT_rich"
repeat_region      /rpt_family="Rf:ATR0062|AF013294 T18A10 repeat 1 from 9736
                    to 10416"
repeat_region      17045..17157
                    /rpt_family="Rf:ATR0062|AF013294 T18A10 repeat 1 from 9736
                    to 10416"
repeat_region      complement(17155..17277)
                    /rpt_family="AT_rich"
repeat_region      17301..17435
                    /rpt_family="Rf:ATR0062|AF013294 T18A10 repeat 1 from 9736
                    to 10416"
repeat_region      17417..17458
                    /rpt_family="AT_rich"
repeat_region      complement(17585..17630)
                    /rpt_family="AT_rich"
repeat_region      complement(17680..17700)
                    /rpt_family="AT_rich"
repeat_region      complement(17826..17846)
                    /rpt_family="AT_rich"

```

```

repeat_region      complement(18263..18314)
                    /rpt_family="(TA)n"
repeat_region      18635..18671
                    /rpt_family="AT_rich"
repeat_region      complement(19023..19043)
                    /rpt_family="(TA)n"
gene               complement(19154..22414)
                    /gene="At2g14700"
                    /note="T6B13.6: predicted by genefinder"
                    /complement(join<19154..19369,19863..19973,20151..20349,
                    22272..>22414))
                    /gene="At2g14700"
                    /complement(join(19154..19369,19863..19973,20151..20349,
                    22272..22414))
                    /gene="At2g14700"
                    /codon_start=1
                    /product="hypothetical protein"
                    /protein_id="AAC69372.1"
                    /db_xref="GI:3810590"
                    /translation="MAKTPVKMADEGLSFVSSPIEKSFVSGTTATLYSGDI
FSLKFGILLVSSDLICLVLRVWEDSRSLSSSTCCQMFEEVVGWGFATLLRL
VRLRVERTLFLSMEAIGRVNISTAMLYIDAVCAKSKTSYGLSYEHSMLRYGV
GGGNMDEVNLGAEEGADQVYKIKELRKALKSLRLIKAEASSSDFELSGKIGVLD
YRPT"
CDS               complement(20454..20480)
                    /rpt_family="AT_rich"
                    repeat_region      21568..21693
                    /rpt_family="AT_rich"
                    repeat_region      21865..21904
                    /rpt_family="AT_rich"
                    repeat_region      complement(21993..22079)
                    /rpt_family="AT_rich"
                    repeat_region      complement(22194..22243)
                    /rpt_family="AT_rich"
                    repeat_region      complement(20454..20480)
                    /rpt_family="AT_rich"
                    repeat_region      21568..21693
                    /rpt_family="AT_rich"
                    repeat_region      21865..21904
                    /rpt_family="AT_rich"
                    repeat_region      complement(21993..22079)
                    /rpt_family="AT_rich"
                    repeat_region      complement(22194..22243)
                    /rpt_family="AT_rich"

Query Match      5.1%; Score 123.2; DB 8; Length 92524;
Best Local Similarity 71.2%; Pred. No. 8.6e-22;
Matches 195; Conservative 0; Mismatches 68; Indels 11; Gaps 2;

QY 271 TAGTTGGTTTATTTATTTATTTGATTTGCTTCTTTATTCGCGGCGAGCGAGTGA 330
    ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 9260 TAGTTTTTTTGGCGATTCACGCTTTGGTTCTCTCGACGCGGCGAC-TTTTACGGC 9202

QY 331 GTTTCGTCAAAACGTTAGCATCCGACTGAGCGTGCAGATGAAGAGAAACAGACG 390
    ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 9201 GATTGATCTCTCCGATTCGATCTGCTCATATTAAGTCAAAAGAACAGACGCA 9142

QY 391 ATTGTGAG-----TAGATGACGCGGCAACGCTGTGTCATCCGCTCGATCGACG 440
    ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 9141 ATTGTGCGCATCCGCTCGATTCAGCGCATAGTGTGTGATCCGCTCATCGACG 9082

QY 441 AGAATACGTTTCGATCCGCTTCGATCCAAATCGAGAGTTTGAGATCTAATCGGAAA 500
    ||||| ||| ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 9081 CGAATAGTGTGCGATCGATCGATTCATTCGAAATCGAGAGTTAAGTTGAAACCGAGA 9022

QY 501 TTGCATTAATCTCATCTCCATCTCTTTGTAAG 534
    ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 9021 ATTATGATCATCTCATCTCCATCTCTTTGTAAG 8988

RESULT 11
AC005957/c      114041 bp      DNA      11near      PLN 11-MAR-2002
LOCUS          Arabidopsis thaliana chromosome 2 clone T15J14 map m1398, complete
DEFINITION
ACCESSION      AC005957
VERSION        AC005957.3      GI:20197491
KEYWORDS
SOURCE
ORGANISM
Arabidopsis thaliana.
Arabidopsis thaliana.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.

```


GENSMAPHLISPELVYDVLGMDWLDHYRVLDCDHGKRVSEPEREGRVLYGVNPTSG
SLVISAQAEKMIKGECEAVYLMISMPESVGOELPSPKSDPFTIELEPGTAPLAKAPY
RMVPEVMAELKLEKLELIGKGFITLNTSPMTSYLPFKKDGSRFLCQDYELRLAPY
KNKPYLPRIDELQDLGATCFESKIDLTSGHOPIAEADVKTAFRTIRGHFEVVA
PFGILNPAAMRLMNSVDFEELDEFVITFIDDLIVKSSPEEDHILRRMEKLEAE
KLFKLKSCSFMORRMGLHIVSEGVSDPEKTEAI RDMRPNTNATEIRHFEVVA
YRFEVGFASMAQPMTKLTKGVDPVSVISDDEFTKSAFMAFOEMDTKQMSATAY
TDSVGLGCVLMORGVIVASROLRKHGECNPTLHDEMAVAFTDHRKSLKIFTPQPE
LNLBRMMELVADYDEIAVHPKSNMVMADLSCKLYGAPGOSVEALYELIALRL
CAVAREPLGLEADVADULTPVRLAQEKBELIASKAESSEIOPAAAGTILVGRVC
VPKDEELRETLSEAHASMFSPHGAATKMDLRYQOWGMRKDVANWAVECQCQL
KVAHQVPGMSLOSLPDMKMDITMDPVVGLSVSRKAIWIVDRLTSAFELAI
RKTDGATVIAKAYSEIVKLGHVSVISDDEFTKSAFMAFOEMDTKQMSATAY
POTGOSEERTIOTLEDMLRMCVLDMGHMAHLSLVYVREAMDVTLWDEGSESTY
GADYVOETTERIVRLKLNKMEADORHRSYADKREPEVGDGVYIKMAALRDPNS
ISEIKSPRYIGPFRIYERGVPAVRIELDPVMAFRKRVHVSMLRKLKHDEDLAK
IPEDQPMTEARLVRLERIKELRRKIPILIKVMDCGVETEETWEPRARKARF
KMPKQVAA
18637. .21635
/gene="T32B20.g"
/join(18637. .20487,20649. .21101,21186. .21635)
/gene="T32B20.g"
/codon_start=1
/evidence=not_experimental
/product="Hypothetical protein T32B20.g"
/protein_id="AAF67364.1"
/db_xref="GI:7682783"
/translation="MATSKDLPSEDDLYEELLRNQSFKLMMRYLIAAESPK
RFTIYERALKAPGSYKWTAYLEREDLRYNLEPHTQSDLSNNTEKRGVTHMKRP
RIWVYIQLTLVQDLITRTFRFALCALPVTQHDRIPEYLVFVSGNIPITSLR
VYRRYLMWDSHIEEFELVKSERMOESARLASVLAEDKFSYIKGKTAKLWLELC
ELLVHANVVISGLNVDAIRGIRKFTDEVDSESLADYIIRKLNLEKADIEEGM
MKVTVRDESVIFDVYSRFEESTVAKKEMMSSEDEENEGNEVDEDEEYNNRL
SVREIQRKINGEFLNDNDYDLRLARLEPIMNRRPALANSVILLRDNHNEHRRY
KIEGNAKQILTYTTEAVRYVDPKAVGKPTLVAVAKLYENKRDVNRVITDKAY
QVNTKIVDLASVWCMAEMELRHRKAGALELMRATVAPTEVRRVRAADENKPE
MKLRSLRMSFYDLEESLQTESTRAVVERKIDLRATPQIIMNVAFLDEKPEY
DAFYVREGVYIFKYPHVKIDIMWYITKFRYRGKTELRARLELHVASPADAV
TLIYQVAKLEEDYGLAKRMKVVEATKRVPEQKLEMYEIIYSRAAIFGVPTREI
YEOAIESGLPHKQVIMCIRKPALESIGEDRARALYKSSQPADRSDPEFNKHH
EFYQHNEDTYRMLKIRKSVASYSOTHTILPENMQKDKLIDVEPAKGLRAGL
PEDEMALEQLSLSTTPTEPAKDGKRGVSAVISOSENGKPKYTGAGEDIELP
DESDDESDDHVEISQKEVPAAVFGGLARKRDEDEGEAGGAOKIGALERIKRK
LNO"
complement(22005. .23540)
/gene="T32B20.e"
complement(join(22005. .22139,23232. .23540))
/gene="T32B20.e"
/codon_start=1
/evidence=not_experimental
/product="Hypothetical protein T32B20.e"
/protein_id="AAF67362.1"
/db_xref="GI:7682781"
/translation="MATSVALTSSPPVSLPLSSSSSFFSNCFVTYTRPNTSLVA
IGRIHQEPTRKPLTNALFGLVLAIVAGVALLFGPKLEIGESICKTVKSFQ
QAAKEFSELKTEPEESVASSQVATSKKEEKTEVESSSENK"
25373. .26596
/gene="T32B20.h"
/join(25373. .25663,26010. .26204,26264. .26596)
/gene="T32B20.h"
/codon_start=1
/evidence=not_experimental
/product="Hypothetical protein T32B20.h"
/protein_id="AAF67365.1"
/db_xref="GI:7682784"
/translation="MKNITFGKIIDVGGKLNKKDCGEREKVDRKIDILFLCKYI
VASKVDGNIDRFLOKIVDVHACETPPWGFETDGMGEGIKVIMNNKCGAKSISI
VEYDVDEKIGFIDLVVDSWRTLRKIKKDKIMWDMHDVRSDDOMEKROHISENY
PHMLREHOKRLITERRQDGNVPPQEKVVAVDVKEVDGENVPPEENYVGNVPP
YEEVAVNEKDVGEVPPPEENVAVDMTEGVEGTGVEIKGVETDVGKCSNMH"
complement(30562. .31125)
/gene="T32B20.d"
complement(30562. .31125)
/gene="T32B20.d"

/gene="T32B20.d"
/codon_start=1
/evidence=not_experimental
/product="Hypothetical protein T32B20.d"
/protein_id="AAF67361.1"
/db_xref="GI:7682780"
/translation="MKNVAKCIRPALHVEVYKRESERVEOISRVGESSVGNVEELN
AFVDNGVAVTIPGELIYGVGNDVEPIETETNVEEMENADSRIFPEANSVEAME
NAVINDRHVDSRDSDEEDNVDRIRIETELSCAGERSHSQKRVILRNGMGGLKY
IKNLPLRRQCKIWLKQISTVLRQOL"
complement(34511. .36010)
/gene="T32B20.c"
complement(join(34511. .34783,34861. .34967,35059. .35134,
35199. .35310,35414. .35504,35744. .36010))
/gene="T32B20.c"
/codon_start=1
/evidence=not_experimental
/product="Hypothetical protein T32B20.c"
/protein_id="AAF67360.1"
/db_xref="GI:7682779"
/translation="MEKVFSDSEITSGNHMSVNGMTSLNRSASPMAPNRFIOESSAA
DGESTRACGVSVSSPPVAVPVDSEIRAFLEKSLUNLACAAYAKRKGFIRQDTSMS
DNGANESQASLASKATPMSSAITSGLSDDEEADDETNNPNTNVRVRLMS
NRESARRRRKQALHSELQLRVENSKLKGLDYQTFENDASVENRYLAKANIEL
RAKYMAEIVYKRLTGFMFMHNPQIYSTVLSLSEISNSPDITTSQVTPPEIISGN
KGLIGCKMNRATSMRRESLEHLQKIRBSVGO"
46929. .49951
/gene="T32B20.i"
/join(46929. .47305,47362. .48144,48829. .49339,49448. .49951)
/gene="T32B20.i"
/codon_start=1
/evidence=not_experimental
/product="Hypothetical protein T32B20.i"
/protein_id="AAF67366.1"
/db_xref="GI:7682785"
/translation="MKMDTEKIDDFAGRLSEVSTRSASLGNTIEMPKLVKFPINS
RKYTHILALAEQVLDVNTLFEEDIVGLKAYEERTCEDEDOGLMLTLKINTQV
GEVNDPFGDEEDVMEIDVLDLKLKLOETQESNGDQOADMALMELEYINENK
QKRLAEENSDKMWYLDNGASNNHMGLEFESKIDKVTCKVRFGDMSRIDDKGS
IPFYSKNDKSLLEDVYITPILRNSIISLGCATPACGVRRKKTITLFDREGELIN
AMRYNRKYKTLVEYKQKQDLSSSDSSMMHDMGIDGNKTYVINKELVIRIK
LIVRQKOTSCILKOVQSPQSTYSASNPMLLVHODLQGITPTPTANKSTGVG
NKGIJDESDEAEKDEEDKNDHEDIOESALDGNVYVIOEIIJMLRSLRSRTPPT
YLKDIILSELEGERLRLAIDEEPMDFESEALKEMRACGEISIKRNTMNVNR
PROAKTIGIKVFKLRNSDGTINKYKARVAKYVOKHGVYDEVEFLFEVYVQPE
GFIYEGKEDQYKILNKALYGLKQAPRANNNHNLNOLMELAKYKSKSESLYRKGENE
LLIYAVYVDLVLVTRSSLRMLLEFRKEVSTKFEPSDGLRLTYIGTEVLQHENGIILS
QERYVNLILEETKDKDCNAVRIIPMCANLKLSKAREASIDEX"
complement(56115. .56312)
/gene="T32B20.b"
complement(56115. .56312)
/gene="T32B20.b"

Query Match 3.9%; Score 93.2; DB 8; Length 81877;
Best Local Similarity 70.2%; Pred. No. 9.8e-14;
Matches 125; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY 100 AAACCGTATACCGTTCGCTCAGAGGTCGCGAACCGGACCGCCGTAACCGGAATC 159
DB 12513 AAACGGAATAACCGATCGGTAGTGTCTGCGAACCGGACCGGACCGTAACCGGAATC 12454

QY 160 CTCGAAGAAGTAATTCGCGATCGGTTCGTAACCAACCTCGGATGCGAGAACCGGAAA 219
DB 12453 TCAATTAAGTTCGCGACCGGATTCCTCTTAATATGTCACCGCGGATTAACCGAAT 12394

QY 220 CTGTCGGTTTTTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCT 277
DB 12393 GTTTCGGTTTTTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCG 12336

RESULT 13
AF002068/c
LOCUS AF002068 54386 bp DNA linear PLN 27-DEC-2000

DEFINITION	Arabidopsis thaliana genomic DNA, chromosome 3, BAC clone: T803.
ACCESSION	AF002068 BA000014
VERSION	AP002068.1 GI:8051676
KEYWORDS	
SOURCE	Arabidopsis thaliana (strain:Columbia) DNA, clone_lib:TAU BAC clone:T803.
ORGANISM	Arabidopsis thaliana
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis. 1 (sites)
AUTHORS	Nakamura,Y.
TITLE	Structural Analysis of Arabidopsis thaliana Chromosome 3. III
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 54386)
AUTHORS	Nakamura,Y.
TITLE	Direct Submission
JOURNAL	Submitted (16-MAY-2000) Yasukazu Nakamura, Kazusa DNA Research Institute, Department of Plant Gene Research, 1532-3, Yana, Kisarazu, Chiba 292-0812, Japan (E-mail:ynakamura@kazusa.or.jp, Tel:81-438-52-3935, Fax:81-438-52-3934)
COMMENT	Address for correspondence: kaos@kazusa.or.jp For the latest information on annotation of this clone, please see http://www.kazusa.or.jp/kaos/cgi-bin/agd/graph.cgi?c=T803 Genes with similarity to proteins in the databases are described in product or 'note' qualifiers. Genes that have no significant protein similarity are described as 'unknown protein'. The software programs used to predict genes include: Grail (Informatics Group, Oak Ridge National Laboratory, http://compbio.ornl.gov/Grail-1.3/), GENSCAN (Chris Burge, MIT, http://CCR-081.mit.edu/GENSCAN.html), NetGene2 (S.M. Hebsgaard, et al., CBS, Technical University of Denmark, http://www.cbs.dtu.dk/services/NetGene2/) and SlicePredictor (Volker Brendel, Stanford University, http://gremmlin.zool.iastate.edu/cgi-bin/sp.cgi). Genes encoding tRNAs are predicted by tRNAscan-SE (Sean Eddy, Washington University School of Medicine, St. Louis, http://genome.wustl.edu/eddy/tRNAscan-SE/). This sequence may not be the entire insert of this clone. It may be shorter because we remove overlaps between neighboring submissions. The 5' clone is F814 and the 3' clone is F1M23.
FEATURES	Location/Qualifiers
source	1..54386
	/organism="Arabidopsis thaliana"
	/strain="Columbia"
	/db_xref="taxon:3702"
	/chromosome="3"
	/clone="T803"
	/clone_lib="TAU BAC"
	/complement(join(827..1073,1360..1604))
CDS	/note="gb AF06086.1
	gene_id:T803.1
	similar to unknown protein"
	/codon_start=1
	/evidence=not_experimental
	/protein_id="BAB01980.1"
	/db_xref="GI:9294129"
	/translation="MTRPIVSGRSFVGTNFSTSTSTYGLORPLVPOISNAQAVI EANDQANIDEANDEQAEQAVIRDLVLRNGAKWYEMDASIEHLVYNVDHLAA TRLKGMVSLAKSGKSEWILSEHWRYMDYWMTPAKAKSEKARSRLFSRDGLGHH CHR"
CDS	/complement(join(2125..2826,2909..2991,3066..3638,4068..5540,5652..6138))
	/note="gene_id:T803.2"
	/pseudo
	/codon_start=1
	/evidence=not_experimental
	/product="transposon protein-like"
CDS	/complement(join(10176..10508,10551..10817,10928..11178,11221..12021,12122..12617,12704..13313,13412..14198))
	/note="gene_id:T803.3"
	/pseudo
	/codon_start=1
	/evidence=not_experimental
	/product="helicase-like protein"
CDS	/complement(join(15838..15949,16034..16176,16277..16393,16512..16827))
	/note="gene_id:T803.4
	unknown protein"
	/pseudo
	/codon_start=1
	/evidence=not_experimental
CDS	/complement(join(17978..18087,18184..18326,18423..18537,18639..18900,18992..19054))
	/note="gb AF34238.1
	gene_id:T803.5
	strong similarity to unknown protein"
	/pseudo
	/codon_start=1
	/evidence=not_experimental
CDS	/complement(join(19768..19837,19928..20012,20096..20204,20385..20434,20533..20587,20666..20803,20987..21109))
	/note="emb CAB87733.1
	gene_id:T803.6
	similar to unknown protein"
	/pseudo
	/codon_start=1
	/evidence=not_experimental
CDS	/complement(join(21775..22104,22193..22405,22491..22614,22735..22934,23012..23141,23223..23310,23406..23533,23626..23722,23887..23952))
	/note="gene_id:T803.7"
	/pseudo
	/codon_start=1
	/evidence=not_experimental
CDS	/product="replication protein A1-like"
	/complement(25983..26370)
	/note="gb AF18641.1
	gene_id:T803.8
	similar to unknown protein"
	/pseudo
	/codon_start=1
	/evidence=not_experimental
CDS	/complement(join(28153..28638,28734..28824,28916..29062))
	/note="contains similarity to retroelement pol
	polyprotein
	gene_id:T803.9"
	/pseudo
	/codon_start=1
	/evidence=not_experimental
CDS	/complement(31714..33511)
	/note="gene_id:T803.10
	unknown protein"
	/pseudo
	/codon_start=1
	/evidence=not_experimental
CDS	35905..37542
	/note="contains similarity to retroelement pol
	polyprotein
	gene_id:T803.11"
	/pseudo
	/codon_start=1
	/evidence=not_experimental
CDS	/complement(join(39525..39821,39951..40199,40294..40432,40572..40790,41263..41644,41747..42020,42161..42415))
	/note="gene_id:T803.12"
	/codon_start=1
	/evidence=not_experimental
	/product="5-epi-aristolochene synthase-like protein;
	terpene synthase"
	/pseudo
	/protein_id="BAB01981.1"
	/db_xref="GI:9294130"
	/translation="MAVARTVFGGLSTYLHQAPLFLKTSQSIFPPPSLSLKPWKHDF VCVAATKTSSTSDLESGRPSILFSPISMGDYLSVSDSEFDDIARETSYMKPYV RDLRLLSSNNSNKKIRLIHLILSLISGYFESRILEMILNKAFAELMILAEEDLLETI SIMEFVFRILYQHKMSCDSPFRKGEGRLEKSLVGVGRMLQLYQAAHIGTSPDYIM

gene
SARGLHYLHTSGSEKGIIRDVKTNTLLDENFVAKMSDFGLSKAGPSMDHTVSTAVK
GSPGYLDPEYFRROOLTREKSDVSGFVGLVAVCAVAVINPTLEKQDINIAEYALMSQ
KORNLSEIIDSINLRGNYSPELSEKYEIGAEKCLADECKNRPMEGVLMSEYQIOME
AMLRKONGENSESSQAVEAPESTFLPACSNODSSETESQTSALHNSA"
complement(17204, .10705)
/gene-"T518.3"
complement(join(7204, .7302,7499, .7549,7870, .7962,
8157, .8232,8340, .8456,8616, .8680,8999, .9112,9212, .9310,
9472, .9563,9891, .10014,10118, .10285,10619, .10705))
/gene-"T518.3"
/note-"Similar to NOB3.3 g1j3880615 putative GTP-binding
protein from C. elegans cosmid gb1z92773. EST gb1AA597331
comes from this gene."
/codon_start=1
/evidence=not experimental
/protein_id="AAD5745.1"
/db_xref="GI:4587514"
/translation="MPKAKAKAGPYERPIILGRSSHKIGIVGLPVGKSTILENTL
TKLSTIPANEPFCTIEPNEARVNPDERFPMLCQYKPKSEIPAFLEIHBAGLVRA
HEGQGLGNELSHIRAVDGIPIHVLRAFEDADIHVDDIVDPVRLLETITBELRLKDE
FVGKIDVDEKSMKRSNDKQLELLOKVKAVLEKGDVRFEDMKTADEILNTD
LSAKPVPVILINERDORRKNKFLPIHAMVOHEGDMIPISGVFEERSIADMRD
EAKYCEENKQSLPRLITKTSAINLIYFTFDPPEVYKQWRORSKAPQAGAIH
TDEKGFCAEMKFEDELKELGNEPVAKAGKTBEGKTYVDDGDIIFRFNVSQGC
K"
complement(11227, .14775)
/gene-"T518.4"
complement(join(11227, .11491,11761, .11945,12058, .12171,
12359, .12446,12540, .12623,12733, .12839,12956, .13024,
13106, .13246,13427, .13516,13598, .13708,13801, .14073,
14128, .14196,14308, .14379,14468, .14551,14659, .14775))
/gene-"T518.4"
/note-"EST gb1z37678 comes from this gene."
/codon_start=1
/evidence=not experimental
/protein_id="AAD5746.1"
/db_xref="GI:4587515"
/translation="MGAEFFPVPFNSNANDMEYTDLVFAVRAKASVONGDTDD
YSQIKVMCLTEADPAVQLETVLSVSVAMIDLVHKKDLLEMSYHSHRS
VMDALVDLISLRITFCVYVSHHEATSGKYLDPGCMVLRNPSQTFKHSQOTL
VKRQEHVPRVHAHKISYILPLAPNLVSIILAOINKIKDKDPSITVYDNLRLRE
NSSIGVYGVSVILMVMERMLDDIVSGCDSDNGMGMELDEAVESTMGEDEFPYG
ALKONTSGVNVSELDKLMVLEFPHLESQNSDRDEVMHOMFELIFSEYNTIN
LYTKTESQFLMFAASLDPENCVGRFSKLDLYLSNTRLMRMSVAAVLASLSHG
KFLPASFVASMRLKRVDECAEYCGTCNDVDPEHMQVAFGCGQALVLELFEMSIYE
IPRQSQFRLESLSHKLNPVLCPLSVSEFEKQAKAGLFISSFLDEEDHEE
LSRAFGFERLDFFPFDPCLLKMSYSISIPNEMFWMVTTTGEDEDELCDDVYIN
GDADSAEPPDDVDESEMTMSTTPKHSMTRETERLLKMRSRIRPSTSPESRILT"
complement(16592, .20157)
/gene-"T518.5"
complement(join(16592, .17013,17112, .17393,17480, .17750,
17836, .18161,18248, .18530,18618, .18690,18769, .18864,
18948, .19345,19448, .19580,19943, .20157))
/gene-"T518.5"
/note-"Strong similarity to gb|U80583 proteinase TWP from
Lycopersicon esculentum and is a member of the Pf|00082
subtilase family."
/codon_start=1
/evidence=not experimental
/protein_id="AAD5747.1"
/db_xref="GI:4587516"
/translation="MDESSLVRFVLLCLVSSVYCLAESDQNAVSSAVIYTLNDR
PSYHFSRESSEDSKHSLSLATSQYRILNLSASLIIRHDSILRNVLKRENYLKLYSH
YLINGFAVILTRKOADRLAREEVENVLFVLPKATTHPOFLGRLMGRLGSGE
YAGEVYVIGLIDGIDPTHPSEFSGDKISGHYVSPHGTGCEVLTIGPPSCNKLIG
ARRHRAESLRGVINSQDASPFDGSRHTHRSVAANGHGPVYVAGHRLGASAM
APRHAIITVAKLRFGFADIIAIDQADQVDILINISIPNRNPSIATFENFI
DMALLSAVKAGIFVVOAGNTGAPAKPSKSEPIVFGATSHDRVSNIIIGNNT
IPGVGLASGRIMKILVLAHALANGTVDALIVGECOSSPFDQKLVQGLICVY
TVREILGSTIKOHLPAKMLTAGLVFYDVSATGOMTSSPDIDGILLISSODOS
ALRLYNSLSLRENGSGKIVGASVAKIVGSRPTVITAPKWAYESARPPDPSDF
VDADIMKPNIVAPGNATMGAMSPGICTNDPOGRMPEEGSTMSAHVYIGIALLIO
KPFHETPAATASALSTYASLDRKEHIMQRTVLPNDISOSPATPDPMGSGFVNATA

gene
ALDPGLIFDIGNEYMKFLGINS SSPVLNVTGESSYSSNLSAASDLNLPSTYAK
LKVRAVLKRWATNATATNETYIVGMADSVSVKPAKFTIGNQOTVLSLVFRA
LMVNSASFRIGLFGDRGVHVNIPVAVIKIAY"
21927, .25768
/gene-"T518.6"
join(21927, .23254,23451, .23642,23780, .23861,23950, .24141,
24253, .24366,24452, .24542,24632, .25305,25421, .25768)
/gene-"T518.6"
/codon_start=1
/evidence=not experimental
/product="T518.6"
/db_xref="GI:4587517"
/translation="MAVTISTNPFVNASLDESNSTFWRLFHQPYNCRRVRLNSR
KLNKVMFCLNLTKEVLOKPGDKGEFSPSDQYLOIMESYAKKKPRLKYE
EDDGGNGNDGVYEVKDKITKSEKIDKFRKRYRSREIYSDRNEVFRNCEIEH
RVATDLNKSNGESSVVALDKSSESVYTPDESPFRKRSKOEHSSPTSGIEG
SKDELDLVEERVORIANKDKSRSSVAVKNSNGESSVTMPKDSFRNRKSK
QEHRRSDTSRGILARGSDELELVEERVORIANKVRNSKDESILVPSDESPRR
GNRKQEWRYQRYSDTSRGIERGSKGDGLDEBERIIEHLANERHIEIRSKLSTRR
IGAKRNDSDSDSLEFAMETPAFRSDESDIDVPATSRVEMERIEKLAVLNGADIN
MPEMOPSKAIRSAKIRYTOYTVNRLIFELKIGMVRVQVTEIMORODKYSKIRI
IYTTALLVILKSRPVBALVPHAMLIQISSYDPMVABSTIATVIGAGAKIKELFYVI
DYMRSPPKKFKFTLEKMDPRLPPDVVYNALVACVQKQKQEGAFVLOQLKQKQ
KPSFVYTGILMEVLAACEKYNLVHEFFRKQKSIIPALAVRVLVNLMKEGSDKAY
HTEVDEMSRGIVGSNGLNGLNARLCSAGRCNGLNKNVNPVYVLLKILNLYKADN
VHTIOFOLKICIRANKPDLVYVYTGLOACVDSGNINKNAYIFDOMKVCSPILYCN
IMLKAYLOGGLFEARELFQKMSDEGDIKNSDEFSEYRPOYTFNIMLDPCEADQK
WDPQVAYRMLHGHVFNKRLHRLRWLEKRSYKGEVMAETHEHMRSRIRPPLI
KERFFRLKEGDHISATSSIALDNGKIEETELAFSTISANSRVLSPREDQSVLRAND
VNRLGSRSSSDSVLGNLILSSCKDYLTKTRHNL"
30747, .33221
gene
Query Match 3.6%; Score 87.4; DB 8; Length 96183;
Best Local Similarity 68.4%; Pred. No. 3.5e-12;
Matches 121; Conservative 0; Mismatches 56; Indels 0; Gaps 0;
Qy 102 AACGATATACCGTTCGATGAGAGTCCGACGACGACCGGTAACCGAAATCT 161
Db 15426 AACGAGAAACCGATGCTATATGCTCCGACGACGACGACGACGAACTCA 15367
Qy 162 CAAAGAATTCGCCGATGCTTGTCTACTGTTCAAACCTCGGTGCCGAGAACGAACT 221
Db 15366 ATTATTAATTTCCGATGCTTGTACCTCTCTCATATGATGGAATCGAAGAAAA 15307
Qy 222 GTCCGTTTTTCGCTTCGCTTCGCTTCCTCCGAACTCCGAGGCTTGTGG 278
Db 15306 ATCGGTTTGTTCGCTTCGCTTGTTCGTTAATCCGAAATCCAGGCTTGTTCG 15250
RESULT 15
AC009526/c 108061 bp DNA linear PLN 07-APR-2000
LOCUS Arabidopsis thaliana chromosome I BAC F2J6 genomic sequence.
DEFINITION complete sequence.
ACCESSION AC009526
VERSION AC009526.4 GI:7523676
KEYWORDS HTG.
SOURCE Arabidopsis thaliana.
ORGANISM Arabidopsis thaliana.
REFERENCE
1 (bases 1 to 108061)
Fedorispiet, N.A., Palm, C.J., Conway, A.B., Conn, L., Hansen, N.F.,
Altati, H., Araujo, R., Huizar, L., Kowley, D., Buehler, E., Dunn, P.,
Gonzalez, A., Krementeskaia, I., Kim, C., Lenz, C., Li, D., Liu, S.,
Luros, S., Schwartz, J., Shinn, P., Totiumi, M., Vystotskaia, V.S.,
Walker, M., Yu, G., Ecker, J., Theologos, A. and Davis, R.W.
Unpublished
2 (bases 1 to 108061)
Fedorispiet, N.A., Palm, C.J., Conway, A.B., Conn, L., Hansen, N.F.,
Altati, H., Araujo, R., Huizar, L., Kowley, D., Buehler, E., Dunn, P.,
Gonzalez, A., Krementeskaia, I., Kim, C., Lenz, C., Li, D., Liu, S.,
Luros, S., Schwartz, J., Shinn, P., Totiumi, M., Vystotskaia, V.S.,
Walker, M., Yu, G., Ecker, J., Theologos, A. and Davis, R.W.
Unpublished
JOURNAL
REFERENCE
AUTHORS

Gonzalez, A., Kremenetskaia, I., Kim, C., Lenz, C., Li, J., Liu, S.,
 Lueros, S., Schwartz, J., Shin, P., Toriumi, M., Vyotskaia, V.,
 Walker, M., Yu, G., Ecker, J., Theologis, A. and Davis, R.W.
 Submitted Submission
 Submitted (26-AUG-1999) DNA Sequencing and Technology Center,
 Stanford University, 855 California Avenue, Palo Alto, CA 94304,
 USA
 3 (bases 1 to 108061)
 Federspiel, N.A., Palm, C.J., Conway, A.B., Conn, L., Hansen, N.F.,
 Altati, H., Araujo, R., Huizar, L., Rowley, D., Buehler, E., Dunn, P.,
 Gonzalez, A., Kremenetskaia, I., Kim, C., Lenz, C., Li, J., Liu, S.,
 Lueros, S., Schwartz, J., Shin, P., Toriumi, M., Vyotskaia, V.,
 Walker, M., Yu, G., Ecker, J., Theologis, A. and Davis, R.W.
 Direct Submission
 Submitted (11-DEC-1999) DNA Sequencing and Technology Center,
 Stanford University, 855 California Avenue, Palo Alto, CA 94304,
 USA
 4 (bases 1 to 108061)
 Federspiel, N.A., Palm, C.J., Conway, A.B., Conn, L., Hansen, N.F.,
 Altati, H., Nguyen, M., Lam, B., Southwick, A., Ecker, J., Theologis, A.
 and Davis, R.W.
 Direct Submission
 Submitted (07-APR-2000) DNA Sequencing and Technology Center,
 Stanford University, 855 California Avenue, Palo Alto, CA 94304,
 USA
 COMMENT
 On Apr 7, 2000 this sequence version replaced gi:6560894.
 Bases 1-4149 of IGF clone F2J6 overlap with bases 91042-95190 of
 TMU clone T10P12 (gb|AC007203) and an unknown number of bases
 overlap at the end of F2J6 overlap with IGF clone F28H19
 (gb|AC006423).
 FEATURES
 source
 1. 108061
 /location="Qualifiers
 /organism="Arabidopsis thaliana"
 /db_xref="taxon:3702"
 /chromosome="1"
 /clone="F2J6"
 complement(4642. .5202)
 /gene="F2J6.1"
 complement(join(4642. .4831,4916. .4978,5051. .5202))
 /gene="F2J6.1"
 /note="Unknown protein; Location of EST F2E8T7, gb|N96384"
 /codon_start=1
 /protein_id="AA63116.1"
 /db_xref="GI:7523677"
 /translation="MAMFVQIVYAGMPLSKYASISGCTPPEVFVROAPALALSP
 FAPFLSSKSSPLSTLLKIFPISLDCGLTSLNLYVALENTTATTAATTNAPISI
 IFVALLEFRITQSLSTLYKPCDDPKVFEVL"
 26596. .28759
 /gene="F2J6.2"
 join(26596. .26754,26977. .27024,27093. .27167,27254. .27315,
 27401. .27467,27581. .27694,27790. .27876,27954. .28024,
 28113. .28178,28273. .28359,28484. .28559,28646. .28759)
 /gene="F2J6.2"
 /codon_start=1
 /product="putative fructose 1,6-bisphosphatase"
 /protein_id="AA63117.1"
 /db_xref="GI:7523678"
 /translation="MDRAADAHRTDMLTTRFVNLNEOSKTPESNGDTILLSHVLCG
 KFYCSAVNKAAGLKLGLAGETNIOGEEKRLDVLNDVFNALVSSGRSVLSEED
 EBAITPEPSRGRKYCVFPLDSSNIDCVSIGTIGIYTLDTDEPTADYLPEN
 ENAAAGACMGSSCMVLSTGTVHGFTLPDSLGEFLLTHDILKIPKNGIYVNEGN
 AOMNDGPTTKYVEKCKRPKDGSPAKSLRYGSMVADVHRTLLYGLIPLYPDKKSPNG
 KRLVYEFVPMSPFLMEAGAGATGKKRALDLYPERKIHNSPIFLGSYDVEEIKALY
 ABERKN"
 30617. .31261
 /gene="F2J6.3"
 join(30617. .30834,30985. .31261)
 /gene="F2J6.3"
 /note="Hypothetical protein"
 /codon_start=1
 /protein_id="AA63123.1"
 /db_xref="GI:7523684"
 /translation="MSDELWNIQHMELGRELFIPYHAYAGAFASNRLSHLGRILNPG

TOSEVARELLEPYOMGLSQVHGHILDRGISHPSYSERTIEETASTLGEVANDNE
 ETTQOIFIRKRVKIDITPEIRFRVRVRESREGAMGNFNEKIQOCTNCCRVNHQY
 SHCP"
 36931. .37350
 /gene="F2J6.4"
 36931. .37350
 /gene="F2J6.4"
 /note="Unknown protein"
 /codon_start=1
 /protein_id="AA63118.1"
 /db_xref="GI:7523679"
 /translation="MADQDEEDLKALKMSMOYNPEPKSKPIEEEGSGSGG
 ESPKARKRLQRELMAAAEKRWLPFKSPSPYKARVLPISTVGGDKDEEVGSCIG
 KELSEESDOLFSTVGFNEVSKSVLAQMTNGIRYDS"
 39054. .41001
 /gene="F2J6.5"
 join(39054. .39133,39234. .39345,39452. .39628,40186. .41001)
 /gene="F2J6.5"
 /note="Hypothetical protein"
 /codon_start=1
 /protein_id="AA63119.1"
 /db_xref="GI:7523680"
 /translation="MGVWTLSDRHHVSDRSSTVEAKTKALVRSCEILPMCGDNN
 RAVIASFTNEEDSSNOKDEAMAGLPIESASDLQILREFTTQASADNKEGTLT
 AFQSRMGALLFLIALSLRGLIEVNLIGCEAVPNVDGMDLGGMFLGISKNDVY
 GFLLSLNCFKCGONLKCPEKPIWYIGSESHYVLEADPVSQENLELRESQIR
 RAPFARQSGGGFTITAEAFQVQVQENIRIPKIKLDICATGFIWSELMQVLELD
 ONLGICIDATGCMKKVPIYHFNGIAKSDINGGQAMAYGGGVPMQRPRTULANS
 VPKWPEETWTCALPPSSSEKDESNQPKVQAHPLVDCITRTWMSRACWSDDPS
 IV"
 42764. .44429
 /gene="F2J6.6"
 join(42764. .43362,43458. .43590,44048. .44132,44221. .44429)
 /gene="F2J6.6"
 /note="Contains bZIP transcription factors basic domain
 protein motif."
 /codon_start=1
 /product="putative transcription factor"
 /protein_id="AA63120.1"
 /db_xref="GI:7523681"
 /translation="MEGGGRGPNOTIIEIHMEAPRORISHRRARSEFFSGESI
 DDLLFPDSIDFSLDLNAPPPPOSOQOPASPMVSVEETSSNGVPPNSLPK
 PEARFGHVRFSFSDFFDLGVTGEKFIATSGEKKRHHHSRSMGEMSSAS
 FNIESILASVSGKSKKMGMDRLAEALLDPKRAKRIILHROSARSKERKIRY
 TGELEKRVQITONATITLSAQVMTLOGTSELTENNHIKRLQALDQOELRALNE
 ALRDELNRKLYACETIPQNGNSNTRQAFSSQASAMQFGKTKQOMSTNGQPSLPST
 MFTKRG"
 complement(44946. .46710)
 /gene="F2J6.7"
 complement(join(44946. .45509,45595. .45731,45811. .46004,
 46088. .46319,46389. .46710))
 /gene="F2J6.7"
 /codon_start=1
 /product="putative histidine decarboxylase"
 /protein_id="AA63121.1"
 /db_xref="GI:7523682"
 /translation="MVGSLSDQFLSMATLIEKLDIISDDPPTAVVTEPPPTNG
 IGAKGGGGGGRRENVLCGRNHTTSLATPEPVNDEPFGDKRAVYASVLAIRKTLVER
 TKNLGLPYNLIDFYGALGOLOHSINNLDGPTIESNIGVHSRFEVGLDWFARLWE
 IERDDYGYITNGCTENGLHDLVAGREPGIILYASSESHYVFKAKARVMRCSEVY
 DTLSDGIDCDLKRKLLANDKDRIILNANQKIVKASVADLDIVTLIEEGCSHR
 FYIHGDGLAFGLAMPYKRAKVFENMPIGSVSGHKGFCGPGCQGLTBMHEITV
 LLSNVEGLASDPAITMGRNGHAPLIMYTLNRRKGYGFOKEQKCLRNHAYLKDRIR
 EAGISAMLNELSTVVERPKDEEFVRWQIACQGDIAHVVMPSVTIEKLNFLKDL
 VKHRLIMVEDSOPPCLASVEGNNCCIPARK"
 complement(47997. .49751)
 /gene="F2J6.8"
 complement(join(47997. .48982,49058. .49751))
 /gene="F2J6.8"
 /note="Hypothetical protein"
 /codon_start=1
 /protein_id="AA63122.1"
 /db_xref="GI:7523683"

and xhoi was ligated to modified lambda FIC-1 vector (Carlucci et al., submitted for publication) digested with BamHI and SalI. This clone is in a modified pluviuscript vector. Please visit our web site (http://www.gsc.riken.go.jp/e/plant/index_e.html) for further details.

FEATURES

Source

Location/Qualifiers
1. 621
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
/clone="RAF109-06-F07"
/clone_lib="RAF19"
/dev_stage="plants at various developmental stages from germination to mature seeds"
/lab_host="DH10B"
/note="Site 1: BamHI; Site 2: SalI; subjected to dehydration (1, 2, 5, 10, 24 hr) and cold (1, 2, 5, 10, 24 hr) treatments"
BASE COUNT 154 a 146 c 147 g 173 t 1 others
ORIGIN

Query Match

Best Local Similarity 100.0%; Score 612; DB 10; Length 621;
Pred. No. 1.8e-139; Mismatches 0; Indels 0; Gaps 0;
Matches 612; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

73 GAAGTGAAGATGAGTGTGTAAGACAAACCGTATACCGTTGGTTCAGAGCTCCG 132
|||||
1 GAAGTGAAGATGAGTGTGTAAGACAAACCGTATACCGTTGGTTCAGAGCTCCG 60
OY AACGACACGACCGCTTAACGAAATCCTCAAAATTCGCGATTCGCTACTGT 192
Db AACGACACGACCGCTTAACGAAATCCTCAAAATTCGCGATTCGCTACTGT 120
OY TCAAAACCTGGTGGCGAAGACCGAACTGCGGTTTTCGGTTCGGGTTTCGGTTT 252
Db TCAAAACCTGGTGGCGAAGACCGAACTGCGGTTTTCGGTTCGGGTTTCGGTTT 180
OY CTCGCAACTCCGAGCGCTACTGTTGTTTATTTTTCACGAGTTTCTCTTTTAT 312
Db CTCGCAACTCCGAGCGCTACTGTTGTTTATTTTTCACGAGTTTCTCTTTTAT 240
OY CGGCGACGACGCGTGTGTTCTGTCAAAAGCTTAACGATCCGCTGAGCGTGCAG 372
Db CGGCGACGACGCGTGTGTTCTGTCAAAAGCTTAACGATCCGCTGAGCGTGCAG 300
OY TAAAGAGAAGAACGACGATTTGTGTAGATCGACGGCGAAGCTGTGATCCCTCG 432
Db TAAAGAGAAGAACGACGATTTGTGTAGATCGACGGCGAAGCTGTGATCCCTCG 360
OY ATCGAGGAGATACGTTTCGATCCGCTTGCATCAATCGGAGAGTTTGAAGATCTTA 492
Db ATCGAGGAGATACGTTTCGATCCGCTTGCATCAATCGGAGAGTTTGAAGATCTTA 420
OY ATCGAAATTCATTAATACATCTCAATCTCTTGAAGAGTCCGAATCCGATCTAC 552
Db ATCGAAATTCATTAATACATCTCAATCTCTTGAAGAGTCCGAATCCGATCTAC 480
OY CACCACTACTGTAACCGCGGTGATTTAGTCCGCAATTCATTCATTCATTC 612
Db CACCACTACTGTAACCGCGGTGATTTAGTCCGCAATTCATTCATTCATTC 540
OY CGGGGATTCGAATGCGAGCTGAGTGAATCTGGGTTTTCAGTGAAGATTTATTTGT 672
Db CGGGGATTCGAATGCGAGCTGAGTGAATCTGGGTTTTCAGTGAAGATTTATTTGT 600
OY CACTCTTTGAAA 684
Db CACTCTTTGAAA 612

RESULT 2
AV548305/c AV548305 556 bp mRNA linear EST 06-SEP-2000
LOCUS
DEFINITION AV548305 Arabidopsis thaliana roots Columbia Arabidopsis thaliana

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
COMMENT

FEATURES

Source
Location/Qualifiers
1. 556
/organism="Arabidopsis thaliana"
/strain="Columbia"
/db_xref="taxon:3702"
/clone="R2L51d11F"
/clone_lib="Arabidopsis thaliana roots Columbia"
/tissue_type="roots"
/note="Vector: pluviuscriptli SK-; Site 1: EcoRI; Site 2: XhoI"
BASE COUNT 171 a 129 c 103 g 153 t
ORIGIN

Query Match

Best Local Similarity 100.0%; Score 556; DB 10; Length 556;
Pred. No. 1e-125; Mismatches 0; Indels 0; Gaps 0;
Matches 556; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY ATGCTACTGAACCAATCAACGAGCCATTAATAAGACTCTTGAGGCAAGAGCTA 1882
Db ATGCTACTGAACCAATCAACGAGCCATTAATAAGACTCTTGAGGCAAGAGCTA 497
OY TGTAGAGATCAGAAATTTCTATAGCTGATGGTGGAGCGGCTTCCGATAGAGC 1942
Db TGTAGAGATCAGAAATTTCTATAGCTGATGGTGGAGCGGCTTCCGATAGAGC 437
OY CTGATCTCAAACTCAACTTTTGTGATCTCAATAGTGTGAAGAGATCTCTCTGCTG 2002
Db CTGATCTCAAACTCAACTTTTGTGATCTCAATAGTGTGAAGAGATCTCTCTGCTG 377
OY GTGTTCTGAGAGCTGTGATGATTCGCAATTTTGCATCACTTTAGGGGATTCGGCA 2062
Db GTGTTCTGAGAGCTGTGATGATTCGCAATTTTGCATCACTTTAGGGGATTCGGCA 317
OY CCAACCTGACCGAGCGATGAGTTCGCAATGTTTGGCTTGTGGTGAAGAGATC 2122
Db CCAACCTGACCGAGCGATGAGTTCGCAATGTTTGGCTTGTGGTGAAGAGATC 257
OY CACATGCGGTTTGGCTTAAGATGATGATCCAGACACATGATTAATCTTCAGGCGTTT 2182
Db CACATGCGGTTTGGCTTAAGATGATGATCCAGACACATGATTAATCTTCAGGCGTTT 197
OY CATCAATTCACCTTGAGTAACACATGTTTCACTGTCATTAATTAAGTTCGCTACCA 2242
Db CATCAATTCACCTTGAGTAACACATGTTTCACTGTCATTAATTAAGTTCGCTACCA 137
OY AGTTGCGTTGAGTATACGTTTTCATATAGACTTGGTGAATTTCTTGGTGAAGC 2302
Db AGTTGCGTTGAGTATACGTTTTCATATAGACTTGGTGAATTTCTTGGTGAAGC 77
OY ATTTTATACCATTTGTAAGGCTTAACTCTGGAATACTGGGGAAATTAATAATAA 2362
Db ATTTTATACCATTTGTAAGGCTTAACTCTGGAATACTGGGGAAATTAATAATAA 17

QY 2363 GTTGATTTCAATCTT 2378
 |||||
 DB 16 GTTGATTTCAATCTT 1

RESULT 3
 AV556975 486 bp mRNA linear EST 06-SEP-2000
 AV556975/c
 LOCUS AV556975 Arabidopsis thaliana green siliques Columbia Arabidopsis
 DEFINITION thaliana cDNA clone S0057109F 3', mRNA sequence.
 ACCESSION AV556975
 VERSION AV556975.1 GI:8728390
 KEYWORDS EST.
 SOURCE thale cress.
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 1 (bases 1 to 486)
 Asamizu, E., Nakamura, Y., Sato, S. and Tabata, S.
 A large scale analysis of cDNA in Arabidopsis thaliana: Generation
 of 12,028 non-redundant expressed sequence tags from normalized and
 size-selected cDNA libraries
 DNA Res. 7, 175-180 (2000)
 20363093
 CONTACT: Erika Asamizu
 The First Laboratory for Plant Gene Research
 Kazusa DNA Research Institute
 Yata 1532-3, Kisarazu, Chiba 292-0812, Japan
 Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/
 Location/Qualifiers

FEATURES
 source 1. 486
 /organism="Arabidopsis thaliana"
 /strain="Columbia"
 /db_xref="taxon:3702"
 /clone="S0057109F"
 /clone_lib="Arabidopsis thaliana green siliques Columbia"
 /tissue_type="green siliques"
 /note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
 XhoI"

BASE COUNT 155 a 114 c 91 g 126 t
 ORIGIN

Query Match 20.3%; Score 486; DB 10; Length 486;
 Best Local Similarity 100.0%; Pred. No. 1.6e-108; Mismatches 0; Indels 0; Gaps 0;
 Matches 486; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1893 CAGAAATTTATGCGTCAGATGGTGAGCGGCTAGCGTTCGATAGAGCTGAATCTCA 1952
 |||||
 DB 486 CAGAAATTTATGCGTCAGATGGTGAGCGGCTAGCGTTCGATAGAGCTGAATCTCA 427
 QY 1953 AACTCACTTTTGATTTACATATGAGTCTGAGAGAGTTTACTTGTGCTCTCGG 2012
 |||||
 DB 426 AACTCACTTTTGATTTACATATGAGTCTGAGAGAGTTTACTTGTGCTCTCGG 367
 QY 2013 AACTGTGATTTGATTCATATTTGCAATCTTATAGGGGATTCGGCACAACACTGAC 2072
 |||||
 DB 366 AACTGTGATTTGATTCATATTTGCAATCTTATAGGGGATTCGGCACAACACTGAC 307
 QY 2073 CCAGCATGAGATTCGCACAAATGTTTGGCTTGTGTGAGAGAGATCCACATGGCGT 2132
 |||||
 DB 306 CCAGCATGAGATTCGCACAAATGTTTGGCTTGTGTGAGAGAGATCCACATGGCGT 247
 QY 2133 TTGCTAGAAATGTTGATTCACGACGACATGATTTACTTACAGCGTTTCATCAATCA 2192
 |||||
 DB 246 TTGCTAGAAATGTTGATTCACGACGACATGATTTACTTACAGCGTTTCATCAATCA 187
 QY 2193 CCTTGAGTAAACAATTTTTCAGTGTCCCAATTTAGTGGGTGACCAAGTTGGTGG 2252
 |||||
 DB 186 CCTTGAGTAAACAATTTTTCAGTGTCCCAATTTAGTGGGTGACCAAGTTGGTGG 127
 QY 2253 AGTAACTGTTTGCATATAGACTGGGTGCTAAATTTCTTGTGTAGCAATTTTATAC 2312

DB 126 AGTAACTGTTTGCATATAGACTGGGTGCTAAATTTCTTGTGTAGCAATTTTATAC 67
 |||||
 QY 2313 CCATTTAGGCTCTTAATCTTGGAAACTTCGGGAAATTAATTAAGTGATTTCA 2372
 |||||
 DB 66 CCATTTAGGCTCTTAATCTTGGAAACTTCGGGAAATTAATTAAGTGATTTCA 7

QY 2373 AATCTT 2378
 |||||
 DB 6 AATCTT 1

RESULT 4
 A1995676 475 bp mRNA linear EST 08-SEP-1999
 A1995676/c
 LOCUS A1995676 A. thaliana, Columbia Col-0, inflorescence-2 Arabidopsis
 DEFINITION thaliana cDNA clone 701516116, mRNA sequence.
 ACCESSION A1995676
 VERSION A1995676.1 GI:5842581
 KEYWORDS EST.
 SOURCE thale cress.
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 1 (bases 1 to 475)
 Chen, J., Momiyama, M., Chan, E., Mooney, M., Carrou, B., Gilliland, D.,
 Wang, X., Hillman, J., Guegler, K., Kim, C., Doyle, M., Broszka, P.,
 Gorgone, G., Burns, D., Griffin, J., Mouanoutoua, M., Nguyen, D., Tan, R.,
 Rose, M., Warren, B., Ton, B., Kasuriy, K., Borillo, C., Carpilo, T.,
 Pollick, J., Suzuki, G., Argentine, C., Shah, S., Nobrigha, A., Murry, L.,
 Turner, C., Krikorian, S., Elder, L. and Hanson, D.
 Arabidopsis thaliana Gene Expression Microarray
 Unpublished (1999)
 CONTACT: David Smoller, Ph.D.
 Genome Systems, Inc., a wholly owned subsidiary of Incyte
 Pharmaceuticals, Inc.
 4633 World Parkway Circle, St. Louis, MO 63134, USA
 Tel: 877-577-2733
 Fax: 314-427-3324
 Email: service@genomesystems.com.

FEATURES
 source 1. 475
 /organism="Arabidopsis thaliana"
 /cultivar="Columbia Col-0"
 /db_xref="taxon:3702"
 /clone="701516116"
 /clone_lib="A. thaliana, Columbia Col-0, inflorescence-2"
 /tissue_type="inflorescence"
 /dev_stage="4 - 7 weeks"
 /note="Vector: pSPORT; Site_1: NotI; Site_2: SalI; cDNA
 library was derived from untreated inflorescence tissue
 from Arabidopsis thaliana, Columbia Col-0, at 4 - 7
 weeks. Plants were grown in 1:1:1 peat
 moss/vermiculite/perlite soil at 22 deg. C +/- 3 deg. C
 under constant light, and watered with fertilizer. cDNA
 synthesis was initiated using a NotI-oligo(dT) primer.
 Double-stranded cDNA was blunt-ended, ligated to SalI adaptors
 , digested with NotI, size-selected, and cloned into the
 NotI and SalI sites of the pSPORT vector."

BASE COUNT 149 a 118 c 90 g 118 t
 ORIGIN

Query Match 19.8%; Score 473.4; DB 9; Length 475;
 Best Local Similarity 99.8%; Pred. No. 2e-105; Mismatches 1; Indels 0; Gaps 0;
 Matches 474; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1874 GAGAACTATGTTGAGATCAGAAATTTCTATGCGTGAATGAGCGGCTAGCGTTC 1933
 |||||
 DB 475 GAGAACTATGTTGAGATCAGAAATTTCTATGCGTGAATGAGCGGCTAGCGTTC 416
 QY 1934 CGATAGAGCTGAATCTCAACTCTTTTGATTTCAATGAGTCTGAAGAGTTTC 1993
 |||||

Db 415 CGATGAGCGCTGATCTCAAACTCAACTTTTGATTTCTACAAATGCTGAGAGAGTTC 356
 QY 1994 TACTGCTGCTGTCTCTGAGAGCTGTGATTTGATTCATTCATTTTGCATTCATTTAGGG 2053
 Db 355 TACTTCTGCTGTCTCTGAGAGCTGTGATTTGATTCATTCATTTTGCATTCATTTAGGG 296
 QY 2054 ATTCCGGCACCACAACTGACCCAGGAGTGTGACACAAATGTTTTGGCTTGTGTGA 2113
 Db 295 ATTCCGGCACCACAACTGACCCAGGAGTGTGACACAAATGTTTTGGCTTGTGTGA 216
 QY 2114 GAGAAATGACATGCGCTTTGGCTTGAAGTGTGATTCACAGACACATGATTAATCT 2173
 Db 235 GAGAAATGACATGCGCTTTGGCTTGAAGTGTGATTCACAGACACATGATTAATCT 176
 QY 2174 CAGGCGTTTCATCATTAATTCACCTTGAATAACAATTTTTCAGTCTCCAAATTAAGGT 2233
 Db 175 CAGGCGTTTCATCATTAATTCACCTTGAATAACAATTTTTCAGTCTCCAAATTAAGGT 116
 QY 2234 GCGTCACCAAGTTGGTGTGATTAATTCATTAATTAAGTGTGGTGTCTAAATTTCTT 2293
 Db 115 GCGTCACCAAGTTGGTGTGATTAATTCATTAATTAAGTGTGGTGTCTAAATTTCTT 56
 QY 2294 GGTGTAGCATTTTTFACCATTTGAGTCTTTTAACTCTTGGAACACTAGCGG 2348
 Db 55 GGTGTAGCATTTTTFACCATTTGAGTCTTTTAACTCTTGGAACACTAGCGG 1

RESULT 5
 AV547260 437 bp mRNA linear EST 07-SEP-2000
 LOCUS AV547260 Arabidopsis thaliana roots Columbia Arabidopsis-thaliana
 DEFINITION CDNA clone RZL28c11f 3', mRNA sequence.
 ACCESSION AV547260
 VERSION AV547260.1 GI:8718674
 KEYWORDS EST.
 SOURCE thale cress.
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 1 (bases 1 to 437)
 Asamizu, E., Nakamura, Y., Sato, S. and Tabata, S.
 A large scale analysis of cDNA in Arabidopsis thaliana: Generation
 of 12,028 non-redundant expressed sequence tags from normalized and
 size-selected cDNA libraries
 DNA Res. 7, 175-180 (2000)

JOURNAL Contact: Erika Asamizu
 MEDLINE The First Laboratory for Plant Gene Research
 COMMENT Kazusa DNA Research Institute
 Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
 Email: asamizu@kazusa.or.jp URL: http://www.kazusa.or.jp/en/plant/.

FEATURES
 source
 1. 437
 /organism="Arabidopsis thaliana"
 /strain="Columbia"
 /db_xref="taxon:3702"
 /clone="RZL28c11f"
 /clone_lib="Arabidopsis thaliana roots Columbia"
 /tissue_type="roots"
 /note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
 XhoI"

BASE COUNT 133 a 95 c 96 g 113 t
 ORIGIN

Query Match 18.2%; Score 437; DB 10; Length 437;
 Best Local Similarity 100.0%; Pred. No. 1.7e-96;
 Matches 437; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 750 GCCAATGACAGGCTGTGTGATGACAAATGACGCTTTTACGCGATTTGTAAGCAAT 809
 Db 1 GCCAATGACAGGCTGTGTGATGACAAATGACGCTTTTACGCGATTTGTAAGCAAT 60

QY 810 CAACGAAGAAGTCACCGCTGAAGATTGGCAGTGAAGAAATGACAGATGTCAAATTAACATC 869
 Db 61 CAACGAAGAAGTCACCGCTGAAGATTGGCAGTGAAGAAATGACAGATGTCAAATTAACATC 120
 QY 870 ACCACAGCTCTGAGAGAAAGCATGTATTAACGTGTACAGTCAATTTTGA-TCCTTAGTC 929
 Db 121 ACCACAGCTCTGAGAGAAAGCATGTATTAACGTGTACAGTCAATTTTGA-TCCTTAGTC 180
 QY 930 TGTGTGCAAGTATTCAGAAACCCCTTTGTAGAGCATGTGATACATGCTATGCTATAGC 989
 Db 181 TGTGTGCAAGTATTCAGAAACCCCTTTGTAGAGCATGTGATACATGCTATGCTATAGC 240
 QY 990 TGTGTGCAATTTGGCAGACCGAGAGACAAAGATCATTTGCACAACTTATTTGCAAG 1049
 Db 241 TGTGTGCAATTTGGCAGACCGAGAGACAAAGATCATTTGCACAACTTATTTGCAAG 300
 QY 1050 TCTGTATTAACAAATTTTAGGCTCCAAATGACTTTTACATATTCGACACAGATTAATC 1109
 Db 301 TCTGTATTAACAAATTTTAGGCTCCAAATGACTTTTACATATTCGACACAGATTAATC 360
 QY 1110 GCGTGGCTTCCATTCAGACACAGAAATGCGTGGTACCTTGCACCGTTTGCATCAATC 1169
 Db 361 GCGTGGCTTCCATTCAGACACAGAAATGCGTGGTACCTTGCACCGTTTGCATCAATC 420
 QY 1170 ATTCAATGCTGGGAGT 1186
 Db 421 ATTCAATGCTGGGAGT 437

RESULT 6
 A1727861 728 bp mRNA linear EST 11-JUN-1999
 LOCUS A1727861
 DEFINITION BNLGH19321 six-day cotton fiber Gossypium hirsutum cDNA 5' similar
 to PHOSPHOHEVALONATE KINASE, mRNA sequence.
 ACCESSION A1727861
 VERSION A1727861.1 GI:5046713
 KEYWORDS EST.
 SOURCE upland cotton.
 ORGANISM Gossypium hirsutum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids II; Malvales; Malvaceae; Gossypium.
 1 (bases 1 to 728)
 Blewitt, M., Matz, E.C., Davy, D.F. and Burr, B.
 ESTs from developing cotton fiber
 Unpublished (1999)
 JOURNAL Contact: Ben Burr
 COMMENT Biology Department
 Brookhaven National Laboratory
 Upton, NY 11973, USA
 Tel: 516-344-3396
 Fax: 516-344-3407
 Email: burr@bnl1.bnl.gov
 Seq primer: T3 primer.

FEATURES
 source
 1. 728
 /organism="Gossypium hirsutum"
 /cultivar="Acala Maxxa"
 /db_xref="taxon:3635"
 /clone_lib="Six-day Cotton fiber"
 /tissue_type="immature fiber"
 /dev_stage="Six days post anthesis"
 /lab_host="XLI-Blue"
 /note="Vector: pBluescript II KS+"

BASE COUNT 213 a 151 c 144 g 219 t 1 others
 ORIGIN

Query Match 17.8%; Score 425.4; DB 9; Length 728;
 Best Local Similarity 76.1%; Pred. No. 1.1e-93;
 Matches 554; Conservative 0; Mismatches 162; Indels 12; Gaps 2;

QY 784 CGGTTTACGCGATTTGAGCAATCAACGAAGTCAAGCTTAAGCTTGGCGATGG 843
 Db 1 CGGTTTACGCGATTTGAGCAATCAACGAAGTCAAGCTTGAAGCTTGGCGATGG 843

Db 1 CGTTTACGCTATGTTAAACCAATTCATGAACCTGGAAGCTGGGAC 60
 QY 844 AATGACAGATGCAATTAACATCCACAGCTCGAGAGAAAGCATGTATTAAC 903
 Db 61 TCTTGACCATGTCAGATCAATCTCCAGCTTCAGAGAAACATGTATTAATTG 120
 QY 904 TCACTGATCATTTGATGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 963
 Db 121 TCTGGAACATTTAACTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
 QY 964 GACATGCGATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020
 Db 181 GAAATGCTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 240
 QY 1021 GAATCATTTGCAACATCTTATTCAGAGCTGCTGCTGCTGCTGCTGCTG 1080
 Db 241 GAGCATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 300
 QY 1081 TTTTACTCATATGGAACACAGATGATGCTGCTGCTGCTGCTGCTGCTG 1140
 Db 301 TTTTACTCATATGGAACACAGATGATGCTGCTGCTGCTGCTGCTGCTG 360
 QY 1141 GATACCTTGCAGCTTGTGATCATATTCATTCATTCATTCATTCATTCAT 1200
 Db 361 GCTACTCTTACACCGCTTATCATATTCATTCATTCATTCATTCATTCAT 420
 QY 1201 TCGAAGCTGAGTAGCAAAACGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1260
 Db 421 TCGAAGCTGAGTAGCAAAACGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
 QY 1261 GTTGACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1315
 Db 481 GTTGACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
 QY 1316 -----AGGAAAGTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1371
 Db 541 CACCAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 600
 QY 1372 CATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1431
 Db 601 CACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
 QY 1432 AGTACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1491
 Db 661 AGTACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
 QY 1492 GGTCTGCC 1499
 Db 721 GGGATGCC 728

RESULT 7
 AV788403/c 424 bp mRNA linear EST 28-MAR-2002
 LOCUS AV788403 RAFL6 Arabidopsis thaliana cDNA clone RAF106-79-G16 3'
 DEFINITION mRNA sequence.
 AV788403
 ACCESSION AV788403.1 GI:19807193
 VERSION AV788403.1 GI:19807193
 KEYWORDS EST.
 SOURCE thale cress.
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliopsida; eudicotyledons; core eudicots;
 Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 1 (bases 1 to 424)
 Seki, M., Narusaka, M., Ishida, J., Kamiya, A., Satou, M., Nakajima, M.,
 Oono, Y., Sakurai, T., Carninci, P., Kawai, J., Itoh, M., Ishii, Y.,
 Arakawa, T., Shibata, K., Shinagawa, A., Muramatsu, M., Hayashizaki, Y.,
 and Shinozaki, K.
 Large scale analysis of Arabidopsis full-length cDNA (2002b)
 TITLE JOURNAL
 COMMENT unpublished (2002)
 Contact: Motoaki Seki
 Plant Functional Genomics Research Group

FEATURES
 Source
 Location/Qualifiers
 1. 424
 /organism="Arabidopsis thaliana"
 /db_xref="taxon:3702"
 /clone="RAFL06-79-G16"
 /dev_stage="plants at various developmental stages from germination to mature seeds"
 /lab_host="DH10B"
 /note="Site 1: Set1; Site 2: XhoI; subjected to dehydration (1, 2, 5, 10, 24 hr) and cold (1, 2, 5, 10, 24 hr) treatments"
 BASE COUNT 138 a 96 c 78 g 112 t
 ORIGIN
 Query Match 17.7%; Score 424; DB 10; Length 424;
 Best Local Similarity 100.0%; Pred. No. 2,7e-93;
 Matches 424; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1953 AACTCAACTTTGGATCTTACATGAGTGTGAAGAGCTTCTGCTGCTGCTG 2012
 Db 424 AACTCAACTTTGGATCTTACATGAGTGTGAAGAGCTTCTGCTGCTGCTG 365
 QY 2013 AGCTGGTGATTTGATGCGATTTTGCATATTCATTTAGGGATTCGGCACAAC 2072
 Db 364 AGCTGGTGATTTGATGCGATTTTGCATATTCATTTAGGGATTCGGCACAAC 305
 QY 2073 CCAGGATGAGTTCGCGCATGTTTGGCTTGTGCTGAGAGATCCAGATGGCGT 2132
 Db 304 CCAGGATGAGTTCGCGCATGTTTGGCTTGTGCTGAGAGATCCAGATGGCGT 245
 QY 2133 TTGCTAGAAAGTGATGTCAGCAACCATGATTTACTTCAGCGCTTCATCAAT 2192
 Db 244 TTGCTAGAAAGTGATGTCAGCAACCATGATTTACTTCAGCGCTTCATCAAT 185
 QY 2193 CCTGAGTAAACATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2252
 Db 184 CCTGAGTAAACATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 125
 QY 2253 AGTATGCTGTTTGCATATGAGTGTGCTAATTTCTGCTGCTGCTGCTGCTG 2312
 Db 124 AGTATGCTGTTTGCATATGAGTGTGCTAATTTCTGCTGCTGCTGCTGCTG 65
 QY 2313 CCATTTGATGCTTTTAATCTTGAAGAACTTGGCGGAAATAAATAAATGAT 2372
 Db 64 CCATTTGATGCTTTTAATCTTGAAGAACTTGGCGGAAATAAATAAATGAT 5
 QY 2373 AATC 2376
 Db 4 AATC 1

RESULT 8
 AV816420/c 424 bp mRNA linear EST 01-APR-2002
 LOCUS AV816420 RAFL9 Arabidopsis thaliana cDNA clone RAF109-90-J06 3'
 DEFINITION mRNA sequence.
 AV816420
 ACCESSION AV816420.1 GI:19858274
 VERSION AV816420.1 GI:19858274
 KEYWORDS EST.
 SOURCE thale cress.
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OY	1799	TTACTTCTGAAAGGGGTGTACATGCTACTGAAACCAATCAACGAACCACTTATTAAG	1858
Db	361	TTAAACCGCAAAAGTGGATGGAGGAGTGACTGAGCTATATAAACGAAAGTTGTCAAAG	420
OY	1859	AACCTTGAAGGCAAGAAACCTATGTGGAGAGATCAGAAATCTTATGCTCAGATGGGTG	1918
Db	421	TATTGCTAAAGGCAAGGAGGTATATGCTTGAATCAGAAATACATCGCAACATGGGAG	480
OY	1919	AGGGGGCTAGCGTTCCGATTAAGCCTGAATCTCAAACTTCAACTTTGGATTCTACATGA	1978
Db	481	AGGAGAGCAATGTGTCCGATAGAACCCGATACACAGACTAAACTTTTGGATCTCTATGA	540
OY	1979	GTCGGAAGAGAGTCTACTTCTGGTGTCTCGAGCGTGGAGATTATAGCCATATTG	2038
Db	541	ATATGGAAGGAGTTTGTGGCCCGAGTTCTCGACAGATGAGATTATGATCAGTCTTTT	600
OY	2039	CAATCACTTTAGGGGATTCGGGACACAATAGACCACAGCATGAGTTCCGACATGTTT	2098
Db	601	CTGTACTCTGGGTGATCTCTAGCAAGAACGTGACAAATATGACATGGAGTTCCGACATGTTT	660
OY	2099	TGGCCTTTTGGTGAAGAAAGATCCACATGCGCGTTTGCTTGAAGTGTGATCCAGCA	2158
Db	661	TGGCCTTTTGTAGTAAAGGAAACCCCTCAAAAGTGTGCTTAAAAAGTGGCATCCACGAT	720
OY	2159	CCACATCTATTACTTCAGGCGTTTCAATCAATTCACCTTGAGTAATAC	2204
Db	721	GCCGGGAATCATGTGTCGACAGTCTTTAGATTAACATCAAGGTGAGC	766

RESULT	10
AV788756/c	
LOCUS	417 bp mRNA linear EST 28-MAR-2002
DEFINITION	AV788756 RABF6 Arabidopsis thaliana cDNA clone RABF06-80-O03 3'
ACCESSION	AV788756
VERSION	AV788756.1 GI:19807546
KEYWORDS	EST.
SOURCE	thale cress.
ORGANISM	Arabidopsis thaliana

REFERENCE	TITLE
1 (Pages 1 to 417)	Large scale analysis of Arabidopsis full-length cDNA (2002p)
SEKI, M., NAKUSKA, M., ISHIDA, J., KAMIYA, A., SATOU, M., NEKAJIMA, M., OONO, Y., SAURU, T., CARINCI, P., KAWAI, J., ITOH, M., ISHII, Y., ARAKAWA, T., SHIBATA, K., SHINAGAWA, A., MORIMATSU, M., HAYASHIZAKI, Y. and SHINOZAKI, K.	
Unpublished (2002)	
Contact: Motoaki Seki	
Comment	

FEATURES
SOURCE
1. .417
Location/Qualifiers
Email: meek@rtc.riken.go.jp
An Arabidopsis full-length cDNA library was constructed, essentially as reported previously (Seki et al., 1998). This clone is in a modified pBluescript vector as a SstI/XhoI insert. Please visit our web site (http://www.gsc.riken.go.jp/e/Plant/index_e.html) for further details.

```

FEATURES
source
location/Qualifiers
1. .417
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
/clone="RAF106-80-003"
/clone_lib="RAF16"
/dev_stage="plants at various developmental stages from germination to mature seeds"
/lab_host="DH10B"
/note="Site_1: SstI; Site_2: XhoI; subjected to dehydration (1, 2, 5, 10, 24 hr) and cold (1, 2, 5, 10, 24

```

	hr) treatments"				
BASE COUNT	136 a	96 c	76 g	109 t	
ORIGIN					
Query Match		16.9%	Score 406;	DB 10;	Length 417;
Best Local Similarity		99.8%;	Pred. No. 7e-89;		
Matches 417;	Conservative	0;	Mismatches	0;	Indels 1; Gaps 1;

Oy	1959	ACTTTTGGATTTCTACAAATGAGTGTGGAAGGAGATTTCTACTGTGTGTTCTCCGGAAGCTGG	2018
Db	417	ACTTTTGGATTTCTACAAATGAGTGTGGAAGGAGATTTCTACTGTGTGTTCTCCGGAAGCTGG	358
Oy	2019	TGGATTTGAATGCCAATATTTTGGCAATCACTTTTAGGSGATTTCCGGCACCAATCTGACCCAGGC	2078
Db	357	TGGATTTGAATGCCAATATTTTGGCAATCACTTTTAGGSGATTTCCGGCACCAATCTGACCCAGGC	298
Oy	2079	ATGGAGTTTCGCACATGTTTTGGCCTTGTGTGTAGAGAAAGATTCACATGGCGTTTGGCT	2138
Db	297	ATGGAGTTTCGCACATGTTTTGGCCTTGTGTGTAGAGAAAGATTCACATGGCGT -TTGGCT	239
Oy	2139	AGAAAGTGGTGATTCACAGCAACACATGTATTTCTTCAGAGCGTTTTCATCAATTCACCTTTGA	2198
Db	238	AGAAAGTGGTGATTCACAGCAACACATGTATTTCTTCAGAGCGTTTTCATCAATTCACCTTTGA	179
Oy	2199	GTAACAACATGTTTTCAGTGTCTCAATTTAATAGTGCCTCACCAAGTTGC-3TTGAGTATA	2258
Db	178	GTAACAACATGTTTTCAGTGTCTCAATTTAATAGTGCCTCACCAAGTTGC3TTGAGTATA	119
Oy	2259	CTGTTTTCGATTTAGACATTGGGTGCTAAATTTCTTGGGTGTAAAGCAATTTTATACCATTG	2318
Db	118	CTGTTTTCGATTTAGACATTGGGTGCTAAATTTCTTGGGTGTAAAGCAATTTTATACCATTG	59
Oy	2319	TAAAGTCTTTAACTCTTGGAAAACTTGGGGGAAAAATAAATTAAGTGTGATTTCAATC	2376
Db	58	TAAAGTCTTTAACTCTTGGAAAACTTGGGGGAAAAATAAATTAAGTGTGATTTCAATC	1

AV796045/c	430 bp	mRNA	linear	EST 29-MAR-2002
LOCUS	AV796045	AV796045	Arabidopsis thaliana	CDNA clone RAFL09-06-F07 3'
DEFINITION	AV796045	RAFL09	Arabidopsis thaliana	CDNA clone RAFL09-06-F07 3'
ACCESSION	AV796045	RAFL09	Arabidopsis thaliana	CDNA clone RAFL09-06-F07 3'
VERSION	AV796045.1	GI:19830028	EST	thale cress.
KEYWORDS				
SOURCE				

REFERENCE
AUTHORS
Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eucotsids II; Brassicales; Brassicaceae; Arabidopsids.
1 (bases 1 to 430)
Seki,M., Narusaka,M., Ishida,J., Kamiya,A., Satoh,M., Nakajima,M.,
Oono,Y., Sakurai,T., Carninci,P., Kawai,J., Itou,M., Ishii,Y.,
Arakawa,T., Shibata,K., Shinagawa,A., Muramatsu,M., Hayashizaki,Y
and Shimozaki,K.
Large scale analysis of Arabidopsis full-length cDNA (2002b)
Unpublished (2002)
Contact: Motoaki Seki

3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
Tel: 81-298-36-4359
Fax: 81-298-36-9060
Email: mseki@rtc.riken.go.jp
An Arabidopsis full-length cDNA library was constructed essentially as reported previously (Seki et al., 1998). cDNA cleaved with BamHI and XhoI was ligated to modified Lambda FLc-1 vector (Carninci et al., submitted for publication) digested with BamHI and SalI. This clone is in a modified pBluescript vector. Please visit our web site (http://www.gsc.riken.go.jp/e/plant/index_e.html) for further details.

FEATURES	Location/Qualifiers
SOURCE	1..430


```

/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
/clone="RAFL09-06-F07"
/clone_lib="RAFL9"
/dev_stage="plants at various developmental stages from
germination to mature seeds"
/lab_host="DH10B"
/note="Site_1: BamHI; Site_2: SalI; subjected to
dehydration (1, 2, 5, 10, 24 hr) and cold (1, 2, 5, 10, 24
hr) treatments"

```

Query Match	16.7%;	Score 401.2;	DB 10;	Length 430;
Best Local Similarity	98.8%;	Pred. No. 1,1e-87;		
Matches 425; Conservative	0;	Mismatches 3;	Indels 2;	Gaps 2;

```
QY      1949 CTCAACTCAACTTTGGATTCTACATGAGTGTGAAGAGTTCATT -GCTGGTGTT 2007
          |||||
Db      430 CTCAACTCAACTTTGGATTCTACATGAGTGTGAAGAGTTCATTGCGTGTT 371
```

[illegible][illegible]

Qy 2128 GGGCTTTGGCTAGAAAGTGGTGATCCAGAACACATGATATTACTTCAGGCGTTTCATCA 2187
|||||
Db 250 GGGCTTTGGCTAGAAAGTGGTGATCCAGAACACATGATATTACTTCAGGCGTTTCATCA 191

Qy 2188 ATTCAACCTTGAGTAAACAACATTGTTCACTGTCCAATTATTA-GGTGCGTCACCAAGTT 2246
|||||
|||
Db 190 ATTCAACCTTGAGTAAACAACATTGTTCACTGTCCAATTATTAAGGTGCGTCACCAAGTT 131

Qy	2247	CGGTTGAGTATAC	TCTGTTTC	CATATAGAC	CTGGTGCTAA	ATTCTTG	GTGAAGCA	TTT	2306
Db	130	CGGTTGAGTATAC <td>CTGTTTC <td>CATATAGAC <td>CTGGTGCTAA <td>ATTCTTG <td>GTGAAGCA <td>TTT</td> <td>71</td> </td></td></td></td></td>	CTGTTTC <td>CATATAGAC <td>CTGGTGCTAA <td>ATTCTTG <td>GTGAAGCA <td>TTT</td> <td>71</td> </td></td></td></td>	CATATAGAC <td>CTGGTGCTAA <td>ATTCTTG <td>GTGAAGCA <td>TTT</td> <td>71</td> </td></td></td>	CTGGTGCTAA <td>ATTCTTG <td>GTGAAGCA <td>TTT</td> <td>71</td> </td></td>	ATTCTTG <td>GTGAAGCA <td>TTT</td> <td>71</td> </td>	GTGAAGCA <td>TTT</td> <td>71</td>	TTT	71

Oy 2307 TTATACCCATTGTAAAGTCTTTACTCTTGCAAACTGCGGGAAAAATAAATAAAGTTTG 2366
|||||
|||
Db 70 TTTATACCATTGTAAAGTCTTTACTCTTGCAAACTGCGGGAAAATAAATAAAGTTTG 11

Qy	2367	ATTCAAAATC	2376
Db	10	ATTCAAAACC	1

RESULT 12	
B0118666/C	
LOCUS	
B0118666	
944 bp	
EBNA	
13002	
FCST 17-APR-2002	

DEFINITION EST6042242 mixed potato tissues Solanum tuberosum cDNA clone STME313
3', end, mRNA sequence.
ACCESSION B011866
VERSION B011866.1 C1.20170639

KEYWORDS	EST.	SOURCE	ORGANISM
Solanum tuberosum	potato.		

Spermatophyta: Magnoliophyta: eudicotyledons: core eudicots; Asteridae: euasterids I; Solanales; Solanaceae; Solanum.
1 (pases 1 to 844)

TITLE	AUTHORS
Restrepo, S., Griffiths, H., van der Hoeven, R., Tsai, J. and Karamycheva, S. A.	Generation of a set of potato cDNA clones for microarray analyses

The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato@tigr.org

This clone is available through the Research Genetics, contact the Research Genetics for further information 1-800-711-6195 or cdna@resgen.com
Seq primer: 77.
Location/Qualifiers
1..844
FEATURES
SOURCE

```

/organism="Solanum tuberosum"
/cultivar="Kennebec Or Binje"
/db_xref="taxon:413"
/clone="STMEC81"
/clone_lib="mixed potato tissues"
/tissue_type="mixed tissues"
/lab_host="corn"

```

```

/rad_josL-50M
/note=Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; supplier: Combination of untreated and Phytophthora
infestans-treated libraries of stolons, leaves, leaflets,
axillary buds of stem explants, petioles, germinating eye-
lubs, or roots."
229 a      215 c      155 g      245 t
BASE COUNT

```

ORIGIN	227	230	233	236
Query Match	15.7%	Score 377.2;	DB 14;	Length 844;
Best Local Similarity	67.58;	Pred. No. 7.6e-82;		
Matches 560; Conservative	0;	Mismatches 268;	Indels 2;	Gaps 2

OY 1316 AAGGAAAGTTGGCGTTCCTCATGTACATGTTATCCATATTGATAGCACAAGAAGCTGCATTT 1375
| | | | |
Db AGGAAAAGAAACATGTAGCTGATCGATGATGATTCATATGATAGCTCAAACTGCATATT 779
| | | | |
OY 1376 GTCTTGACACAAGGAAGTCCGGAAGTGGCTTTGATCTCACGCTGCTGTTCTATGGAAGTC 1435

Db 778 GCATTGCACAGGGTAAAGTTGGCAGTGGSTTTGATGTTTAGTCTGCAGTTATGGAAGTC 719

Qy 1436 AGCGTTATGTCGCTCTCTCCAGAGTCTTGTCATTTCCGAGGTTGCAGTAACAGTC 1495

Db 718 AACGTTAATCCGGTTTTCGCTGAGAGTCTCTTTCTCCGACGAATGCAGGTAATGGCAA 659

OY 1496 TGCACATTA - ATGAAGTATTGGTCAATTTTGAAAGGAAAATGGACATTAAGAACAAT 1554
| | | | |
Db 658 CACCCTAATAGGAGTAGTATTGATGTCCTTAAAGCAAAGTGGCACCATGAGAGACC 599
| | | | |
OY 1555 GAGTCTCTTTACCAACACTGATGAATCTTTTCCTTGGAGAACCTGGAAGTGGTGATCC 1614
| | | | |

Db 598 AAGTTTCATTGCGCTTCCTTTAATGACATATTACTTGGAACACGGGCTGTGGAGGATCT 539

Qy 1615 TCCACA -CCATCAATGTAAGTGCAGTAAGAAGTGGCAAAATGTCGTGATCCAGAAAGGC 1673

Db 538 TCACACCCATCATGTTGGAGCTGTTAAGAATAAGGACAGAGGTCCGACCTCCAGGATTC 479

Accession	Sequence	Position
QY 1674	ACGGAAGAAACATGGCAGAAATTTTTCAGATCTCAAAATTTAGAACTGGGAATCTTAAGCTTAACCA	1733
Db 478	TT	419
QY 1734	TCTGACCAATTTAGCTAAAGACCACTGGGATGTTTATCTACAGATCATTTAACTCTTTAG	1793
Db 418	CTTTTCTTTTATTTTCTTCTGAGAGATCATTTTAACTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT	359

[illegible]

Accession	Sequence	Length
QY 1914	GGGTGAGGGCGGTACGGTTCGATGAGCGTGAATCTCAACTCAACTTTGGATTCTAC	1973
Db 298	TAAAGAAATTAATTAGAGGCTCGAGAGTCTATCGTTCGGATTAGTATCAGATGCGCAAGAT	239
QY 238	GGGGAACCTCTGCGAATTCCTCAATGAGAACCAAAATGCGCAATCTCAGCTCTGATGCTC	179

[illegible]

Db 567 CACCTTTTCTCGATCACTTTAAATGAGAGAAATCAACTGAGAAAATTGCAAGCCTG 626
OY 1211 AAGTAGCAAAACCTGG-CTTAGGTTCTCTGAGCAATGACAACAGCTGTGAGCT 1269
Db 627 AAGTTGCAAGACTGGAATTAGGGTCGTCAGCAGCTATGACAAGTGTGCGGCC 686
OY 1270 CTGTTACATTATCTTG 1286
Db 687 TTGCTTCATTATCTTTG 703

Search completed: May 3, 2003, 19:54:21
Job time : 2265 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 Compen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 3, 2003, 17:59:35 ; Search time 716 Seconds
(without alignments)
15719.537 Million cell updates/sec

Title: US-09-988-863a-1
Perfect score: 2396
Sequence: 1 gtcgaccacgcgtcgcggc.....ttctcaaaaaaaaaaaaaa 2396

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 6104442 seqs, 2348745460 residues
Total number of hits satisfying chosen parameters: 12208884

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Pending_Patents_NA_New:*
1: /cgn2_6/pdata/1/pna/US06_NEW_COMB.seq:*
2: /cgn2_6/pdata/1/pna/US06_NEW_COMB.seq:*
3: /cgn2_6/pdata/1/pna/US07_NEW_COMB.seq:*
4: /cgn2_6/pdata/1/pna/US08_NEW_COMB.seq:*
5: /cgn2_6/pdata/1/pna/US09_NEW_COMB.seq:*
6: /cgn2_6/pdata/1/pna/US09_NEW_COMB.seq:*
7: /cgn2_6/pdata/1/pna/US09_NEW_COMB.seq:*
8: /cgn2_6/pdata/1/pna/US10_NEW_COMB.seq:*
9: /cgn2_6/pdata/1/pna/US10_NEW_COMB.seq:*
10: /cgn2_6/pdata/1/pna/US60_NEW_COMB.seq:*
11: /cgn2_6/pdata/1/pna/US60_NEW_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	229.2	9.6	527	11	US-09-909-745A-21 Sequence 21, Appl
2	194.4	8.1	527	6	US-60-288-292-33093 Sequence 33093, A
3	171.8	7.2	379	8	US-10-389-048-18322 Sequence 18322, A
4	158.2	6.6	675	7	US-09-837-604A-65010 Sequence 65010, A
5	150.4	6.3	539	6	US-09-909-745A-19 Sequence 19, Appl
6	140	5.8	388	7	US-09-837-604A-34763 Sequence 34763, A
7	135.8	5.7	585	7	US-09-837-604A-38723 Sequence 38723, A
8	113.8	4.7	296	6	US-09-615-606A-54496 Sequence 54496, A
9	107.2	4.5	249	6	US-09-909-745A-15 Sequence 15, Appl
10	107.2	4.5	249	6	US-09-909-745A-17 Sequence 17, Appl
11	73.6	3.1	1269	8	US-10-369-493-25849 Sequence 25849, A
12	57.6	2.4	1356	8	US-10-369-493-25596 Sequence 25596, A
13	43.8	1.8	1506	1	PCR-US02-40225-2218 Sequence 2218, Ap
14	43.8	1.8	1506	9	US-10-320-797-2218 Sequence 2218, Ap
15	42	1.8	256324	7	US-09-947-911-95 Sequence 95, Appl
16	41.8	1.7	1664976	1	PCR-US02-40225-1218 Sequence 1218, Ap
17	41.4	1.7	1816	9	US-10-320-797-1218 Sequence 1218, Ap
18	41.4	1.7	1816	9	PCR-US02-40225-218 Sequence 218, App
19	41.4	1.7	3156	9	US-10-320-797-218 Sequence 218, App
20	41.4	1.7	3156	9	US-10-320-797-218 Sequence 218, App
21	41	1.7	360	7	US-09-837-604A-18934 Sequence 18934, A
22	40.4	1.7	1206	8	US-10-369-493-27427 Sequence 27427, A

23	40.4	1.7	3468	9	US-10-282-122A-10932 Sequence 10932, A
24	39.6	1.7	1536	8	US-10-369-493-46489 Sequence 46489, A
25	39.4	1.6	455	8	US-10-391-339-5507 Sequence 5507, Ap
26	38.8	1.6	648	7	US-09-675-784A-5811 Sequence 5811, Ap
27	38.8	1.6	1141	6	US-09-806-708B-22 Sequence 22, Appl
28	38.4	1.6	247	7	US-09-837-604A-56235 Sequence 56235, A
29	38.4	1.6	527	7	US-09-837-604A-5480 Sequence 5480, Ap
30	38	1.6	405	7	US-09-837-604A-23709 Sequence 23709, A
31	37.8	1.6	264665	7	US-09-949-016-13747 Sequence 13747, A
32	37.6	1.6	353	9	US-09-837-604A-17362 Sequence 17362, A
33	37.4	1.6	9652	9	US-10-311-455-882 Sequence 882, App
34	37.4	1.6	135030	7	US-09-949-016-14896 Sequence 14896, A
35	37.4	1.6	3673778	9	US-10-312-841-2 Sequence 2, Appl
36	37.2	1.6	287	7	US-09-837-604A-464 Sequence 464, Appl
37	37.2	1.6	2209	1	PCR-US02-41115-94 Sequence 94, Appl
38	37.2	1.6	17211	9	US-10-311-455-628 Sequence 628, App
39	37.2	1.6	50563	7	US-09-949-016-15821 Sequence 15821, A
40	37	1.5	411	6	US-09-615-606A-43691 Sequence 43691, A
41	37	1.5	540438	7	US-09-947-911-322 Sequence 322, App
42	36.8	1.5	274	7	US-09-837-604A-10493 Sequence 10493, A
43	36.8	1.5	546	7	US-09-837-604A-64175 Sequence 64175, A
44	36.8	1.5	601	7	US-09-949-016-166383 Sequence 166383, A
45	36.8	1.5	6264	9	US-10-240-452-67 Sequence 67, Appl

ALIGNMENTS

RESULT 1	US-09-909-745A-21	
;	Sequence 21, Application US/09909745A	
;	GENERAL INFORMATION:	
;	APPLICANT: Falco, S. Carl	
;	APPLICANT: Ramodu, Omolayo O.	
;	TITLE OF INVENTION: Squalene Synthesis Enzymes	
;	FILE REFERENCE: B1112 US CIP	
;	CURRENT APPLICATION NUMBER: US/09/909,745A	
;	CURRENT FILING DATE: 2001-07-20	
;	PRIOR APPLICATION NUMBER: 09/433,242	
;	PRIOR FILING DATE: 1999-11-04	
;	PRIOR APPLICATION NUMBER: 60/107,241	
;	PRIOR FILING DATE: 1998-11-05	
;	NUMBER OF SEQ ID NOS: 26	
;	SOFTWARE: Microsoft Office 97	
;	SEQ ID NO 21	
;	LENGTH: 757	
;	TYPE: DNA	
;	ORGANISM: Oryza sativa	
;	US-09-909-745A-21	
Query Match	9.6%; Score 229.2; DB 6; Length 757;	
Best Local Similarity	61.3%; Pred. No 2.3e-57;	
Matches 369; Conservative	0; Mismatches 233; Indels 0; Gaps 0;	
Oy	1601 GAACTGGTGATCCTCCACACCATCAATGTAGTGCAGTAAGAGAGCGCAATGCTG 1660	
Db	9 GTATCGAGAGATCATCCACATCCATCAATGTTGATCTGCAACAGCGCAGAGCTAG 68	
Oy	1661 ATCAGAGAGAGCGCAGGAAATTCGAGATTTGTAGATGCAATTTAGAACTGAAA 1720	
Db	69 ACCCTCAGAAATCCAGAGAGAGATTAATTTGGGATTTGTAATTCAGTCTTGA 128	
Oy	1721 CTAGCTAAAGAGATCTGAGCAATTTACTTAAGACCACTGGAGGTTTATCTACGATCA 1780	
Db	129 ACCACTGAGAGACCTTAACCAATCTTCTGAGATCTAGGAGAGCTTATGATCTTTT 188	
Oy	1781 TTAAGCTTTAGTGTCTTACTTCTGAAAGATGGGTGTACAGCTACTGAAACCAATCA 1840	
Db	189 TAGATCTCTGATCGCTCCTACCGTGAAGTGAAGAGAGGTGCTACCAATCAATCAATC 248	
Oy	1841 ACAGAGCATTTATTAAGAACTCTTAGAGCGCAGAGAGCTATGTTAGATCAAGAAATTC 1900	
Db	249 AAGAACTAATTTGTTAGATCAATTACTGCGCGAAGAGATGCTTTCTTGAATATAGGCTTC 308	

;; CURRENT APPLICATION NUMBER: US/09/837, 604A
;; CURRENT FILING DATE: 2001-04-18
;; PRIOR APPLICATION NUMBER: US 60/197, 872
;; PRIOR FILING DATE: 2000-04-19
;; NUMBER OF SEQ ID NOS: 81288
;; SEQ ID NO 65010
;; LENGTH: 675
;; TYPE: DNA
;; ORGANISM: Oryza sativa nipponbare
;; FEATURE:
;; OTHER INFORMATION: Clone ID: LIB3598-023-P1-K6-D6
;; US-09-837-604A-65010

Query Match 6.6%; Score 158.2; DB 7; Length 675;
Best Local Similarity 60.3%; Pred. No. 7.4e-33;
Matches 279; Conservative 0; Mismatches 183; Indels 1; Gaps 1;

Qy 1741 AATTAAGCTAAAGACCACTGGGATGTTATCTACGATCATTAAGTCTTGTAGTGTCTT 1800
Db 1 AACTTGTGTAAAGACCACTGGGATGTTATCTACGATCATTAAGTCTTGTAGTGTCTC 60
Qy 1801 AC-TTCTGAAAAGTGGGTGTTACATGCTACTGAAACCAATCAAGAGCATTAATAAGA 1859
Db 61 ACCTGCAATTAAGGTGACAGAGAGGTGGCTACCAATCAATCAAGAGCATTAATTGTAGATC 120
Qy 1860 ACTCTTAAGAGCAAGAGAGATGTTAGGATTCAGAAATTTCTTATGCGTCAAGATGGTGA 1919
Db 121 ATTACTGGCCGCAAGAGATGTTCTTCTTGAATAATAGGCTTCTATGCGAGATGGCAT 180
Qy 1920 GCGGCTTACGCTCCGATAGAGCTGAATCTCAAACTCACTTTTGTATCTACATAGAG 1979
Db 181 AGCACTGTGTCTCCAAATGAGCCAGAAATCAAACTCAACTTTCAGTGCACATATGAA 240
Qy 1980 TGCTGAGAGATGTTACTTGTGTGTTCTGCTGCTGAGTGTGATGATTCATATTTGC 2039
Db 241 TATGAGGGGTGTTCTACTAGCTGAGTTCCTGGGCGCGTGTGATGAGTGTCTC 300
Qy 2040 AATCACTTTAGGGATTCGCGACCAAACTGACCCAGCATGATTCGACATATGTTT 2099
Db 301 AGTAGTTTGGGTGAAGAGATGATGCTGTAGCCAAAGCTTGAGACTGAGTGTCTC 360
Qy 2100 GCGCTTGTGTGAGAGAGATCCACATGCGCTTGGCTAGAAAGTGTGATCCAGAAC 2159
Db 361 CCCCTTCTTGTGAGAGATCCCGGAGGTGTTCATTGGAAGGTGGACCAAGAAC 420
Qy 2160 CACATGATTAATCTACGAGCGTTTCTCATTCATTCACCTTGAGTAA 2202
Db 421 AAGGAGGTGTCACCGCTGATATCATGATCAATAAATGA 463

RESULT 5

US-09-909-745A-19
;; Sequence 19, Application US/09909745A
;; GENERAL INFORMATION:
;; APPLICANT: Falco, S. Carl
;; APPLICANT: Farnodu, Omolayo O.
;; TITLE OF INVENTION: Squalene Synthesis Enzymes
;; FILE REFERENCE: B1112 US CIP
;; CURRENT APPLICATION NUMBER: US/09/909, 745A
;; CURRENT FILING DATE: 2001-07-20
;; PRIOR APPLICATION NUMBER: 09/433, 242
;; PRIOR FILING DATE: 1999-11-04
;; PRIOR APPLICATION NUMBER: 60/107, 241
;; PRIOR FILING DATE: 1998-11-05
;; NUMBER OF SEQ ID NOS: 26
;; SOFTWARE: Microsoft Office 97
;; SEQ ID NO 19
;; LENGTH: 539
;; TYPE: DNA
;; ORGANISM: Oryza sativa
;; FEATURE:
;; NAME/KEY: unsure
;; LOCATION: (287)

;; OTHER INFORMATION: n=A, C, G, or T
;; FEATURE:
;; NAME/KEY: unsure
;; LOCATION: (440)
;; OTHER INFORMATION: n=A, C, G, or T
;; FEATURE:
;; NAME/KEY: unsure
;; LOCATION: (454)
;; OTHER INFORMATION: n=A, C, G, or T
;; FEATURE:
;; NAME/KEY: unsure
;; LOCATION: (460)
;; OTHER INFORMATION: n=A, C, G, or T
;; FEATURE:
;; NAME/KEY: unsure
;; LOCATION: (465)
;; OTHER INFORMATION: n=A, C, G, or T
;; FEATURE:
;; NAME/KEY: unsure
;; LOCATION: (486)
;; OTHER INFORMATION: n=A, C, G, or T
;; FEATURE:
;; NAME/KEY: unsure
;; LOCATION: (488)
;; OTHER INFORMATION: n=A, C, G, or T
;; FEATURE:
;; NAME/KEY: unsure
;; LOCATION: (495)
;; OTHER INFORMATION: n=A, C, G, or T
;; FEATURE:
;; NAME/KEY: unsure
;; LOCATION: (530)
;; OTHER INFORMATION: n=A, C, G, or T
;; FEATURE:
;; NAME/KEY: unsure
;; LOCATION: (533)
;; OTHER INFORMATION: n=A, C, G, or T
;; US-09-909-745A-19

Query Match 6.3%; Score 150.4; DB 6; Length 539;
Best Local Similarity 62.6%; Pred. No. 9.6e-31;
Matches 266; Conservative 0; Mismatches 157; Indels 2; Gaps 2;

Qy 1601 GAAGGTGATGCTTCCACACCATCAATGATGATGAGTAAAGAGTGAATGCTG 1660
Db 2 GATGCGAGATGATCCATCCATCAATGATGATGATGATGATGATGATGATGATGATG 61
Qy 1661 ATCCAGAGAGGACGAGAGAAATGCGAGATTCAGATCCAAATTTAGAACTGAA 1720
Db 62 ACCCTCAGAAATCCAGAGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAG 121
Qy 1721 CTAAGCTAAAGATGATGAGCAATTTAGTAAAGACACATGAGATGTTATCTACAGTCA 1780
Db 122 ACCAATGAGAGAACTTAACAACTTGTGAAGATCACTGGGAGAACCTATGATGTTT 181
Qy 1781 TTAAGCTTGTGATGCTTACTTCTGAAGAGTGGGTGTTACATGCTACTGAACCATCA 1840
Db 182 TAGCATCTGTAGTGTCTCACAGTCAATGATGAGAGAGTGGCTTACCAATCAACATC 241
Qy 1841 ACGAAGCCATTTAAGAACTCTTAGAGGCAAGAGATGTTGTTAGAGGATCAGAAATTC 1900
Db 242 AAGAACTAATTTAGATCATTAAGTGGCGGCAAGAGATGCTTCCNTGAATAAGGCTTC 301
Qy 1901 TTATG-CGTAGATGAGTGGAGGCGCTACGCTTCCGATAGAGCTGATTCAAATCA 1959
Db 302 ATATGCCAAGAGATGAGATGAGTGGTTCATTAATGAGCCAGAAATCAAACTCA 361
Qy 1960 CTTTGGATTTAGCAATGAGTGTGAGAGAGTGTCTACTGCTGCTGCTGCTGAGCTGCT 2019
Db 362 CTTCTGATGCCATTAATGATGAGAGGCTTCTACTACTGAG-ATTCTTGAGGCGGCT 420
Qy 2020 GGATT 2024
11 11


```

Oy 1057 ATACAAATATTAGGCTCCAAATGACTTTTACTCATATCGGAACCAATAGAAATGGCTGGG 1116
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3 ATTACAAATTTTGGTGCAGAGATTTTCAATGCTATCGCAATGACATCAGAGAGCTGA 62
Oy 1117 CTTCATTGACACCAAGATCGGTGGGTACCTTGCAACGTTTGCAATCATTCATCAAT 1176
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 63 CTCCCTTTGACCTCCCGAATGCAATGCGCAACCTTTCCTGCTTTACTCCATTCCTTCAAT 122
Oy 1177 GCTGGAGAGTCAAATGCTGCTAATTCACAGCTGGAAGTGAAGAAAACGCTTAGGTTCT 1236
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 123 ACCGATGATGCCGATGAGACGGAATTCGAAGCTGGAAGTGGCTAAACCTGGTTGGGCTCT 182
Oy 1237 TCTGAGCAATGACAAACAGCTGTGGTGCAGCTCTGTATACATTATCTTGAGTGTGAC 1296
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 183 TCTGAGCGAATGCAATGAGCTGTAGTGTGCTCTACTGCAATTAACCTGATGCTGTCAAG 242
Oy 1297 CTATC 1301
    ||| |||
Db 243 CTTTC 247

```

```

RESULT 9
US-09-909-745A-15/c
; Sequence 15, Application US/09909745A
; GENERAL INFORMATION:
; APPLICANT: Falco, S. Carl
; APPLICANT: Famodu, Omolayo O.
; TITLE OF INVENTION: Squalene Synthesis Enzymes
; FILE REFERENCE: B1112 US CIP
; CURRENT APPLICATION NUMBER: US/09/909,745A
; PRIOR FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: 09/433,242
; PRIOR FILING DATE: 1999-11-04
; PRIOR APPLICATION NUMBER: 60/107,241
; PRIOR FILING DATE: 1998-11-05
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 15
; LENGTH: 249
; TYPE: DNA
; ORGANISM: Zea mays
US-09-909-745A-15

```

```

Query Match 4.5%; Score 107.2; DB 6; Length 249;
Best Local Similarity 64.5%; Pred. No. 5.7e-19;
Matches 160; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

```

```

Oy 1789 TGTAGTGTCTTACTTCTGAAAAGTGGGTTTACATGCTACTGAAACCAATCAACGAAGCC 1848
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 249 TGTAGTGTCTTACATATGGAAGTGGGAGAGAGTGGCTACTAACCAATCAACGAATTA 190
Oy 1849 ATTATTAAGAAGCTTTAGAGCAAGAGAGTATGTAGATCAGATTTCTTATGCGT 1908
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 189 ATTATTAAGATCTTCTTGGCTGCAAGGAGTGTCTTGAATTAAGGCTTACACATGCA 130
Oy 1909 CAGATGGTGAAGCGGCTAGCTCCGATAGAGACCTGAATCTCAACCAATCAATTTGGAT 1968
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 129 GAGATGGGATGACAGCTGTGCTTCCAAATGAGCCAGATTCACAAACAGGCTACTAGAT 70
Oy 1969 TCTACAAATGAGTGTGAAGAGAGTCTACTTCTGCTGTGCTGTGAGCTGTGAGATTGAT 2028
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 69 GCCACTATGAATATGAGAGGTGTTTACTGCTGAGATTCGGGAGAGCTGTGCTTCGAT 10
Oy 2029 GCCATATT 2036
    ||| |||
Db 9 GCAGTGT 2

```

```

RESULT 10
US-09-909-745A-17
; Sequence 17, Application US/09909745A
; GENERAL INFORMATION:
; APPLICANT: Falco, S. Carl

```

```

; APPLICANT: Famodu, Omolayo O.
; TITLE OF INVENTION: Squalene Synthesis Enzymes
; FILE REFERENCE: B1112 US CIP
; CURRENT APPLICATION NUMBER: US/09/909,745A
; CURRENT FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: 09/433,242
; PRIOR FILING DATE: 1999-11-04
; PRIOR APPLICATION NUMBER: 60/107,241
; PRIOR FILING DATE: 1998-11-05
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 17
; LENGTH: 249
; TYPE: DNA
; ORGANISM: Zea mays
US-09-909-745A-17

```

```

Query Match 4.5%; Score 107.2; DB 6; Length 249;
Best Local Similarity 64.5%; Pred. No. 5.7e-19;
Matches 160; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

```

```

Oy 1789 TGTAGTGTCTTACTTCTGAAAAGTGGGTTTACATGCTACTGAAACCAATCAACGAAGCC 1848
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1 TGTAGTGTCTTACATATGGAAGTGGGAGAGTGGCTACTAACCAATCAACGAATTA 60
Oy 1849 ATTATTAAGAAGCTTTAGAGCAAGAGAGTATGTAGATGATGAGTCAATTCCTTATGCGT 1908
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 61 ATTATTAAGATCTTGTGCTGCAAGGATGCTGCTGCTGAGATTAAGGCTTACATGCA 120
Oy 1909 CAGATGGTGAAGCGGCTAGCTCCGATAGAGACCTGAATCTCAACCAATCAATTTGGAT 1968
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 121 GAGATGGGATGACAGCTGTGCTTCCAAATGAGCCAGATTCACAAACAGGCTACTAGAT 180
Oy 1969 TCTACAAATGAGTGTGAAGAGAGTCTACTTCTGCTGTGCTTCCCTGAGAGCTGTGATTTGAT 2028
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 181 GCCACTATGAATATGAGAGGTGTTTACGCTGTGAGGTTCCGGAGAGCTGTGCTTCGAT 240
Oy 2029 GCCATATT 2036
    ||| |||
Db 241 GCAGTCTT 248

```

```

RESULT 11
US-10-369-493-25849
; Sequence 25849, Application US/10369493
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 25849
; LENGTH: 1269
; TYPE: DNA
; ORGANISM: Schizosaccharomyces pombe
US-10-369-493-25849

```

```

Query Match 3.1%; Score 73.6; DB 8; Length 1269;
Best Local Similarity 48.4%; Pred. No. 1.6e-09;
Matches 239; Conservative 0; Mismatches 249; Indels 6; Gaps 1;

```

```

Oy 1209 TGAAGTACGAAACAGCTAGTGTCTTCTGAGCAATGACAAACAGCTGTGGTTGCAGC 1268
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 438 TCAAGTACATTAAGACTGTTTGAAGAGCTGGCAGCTATGATTAACAAAGTCTATAGTTTC 497

```


THIS PAGE BLANK (uspto)


```

|||||
Db 1381 GCACAGGAGGAGGAGGAGGAGTGGTTCACAGTGGCTGTATGAGAGTACGCT 1440
Oy 1441 TATGTCGCTTCGACAGAGCTTGTATTTGTCAGAGTGGACGTACAGTGGCA 1500
Db 1441 TATGTCGCTTCGACAGAGCTTGTATTTGTCAGAGTGGACGTACAGTGGCA 1500
Oy 1501 TTAATGAAGTATTTGTAACATTTTGAAGGAAATGGGACATTAAGAGCTGAGTTC 1560
Db 1501 TTAATGAAGTATTTGTAACATTTTGAAGGAAATGGGACATTAAGAGCTGAGTTC 1560
Oy 1561 TCTTACCAACACATGATGATTTTCTTGGAGAACTGGAGTGGATCTCCACA 1620
Db 1561 TCTTACCAACACATGATGATTTTCTTGGAGAACTGGAGTGGATCTCCACA 1620
Oy 1621 CCATCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1680
Db 1621 CCATCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1680
Oy 1681 AACTGGCAGAAATTTGTCAGATGCAAAATTTGAAGTGAAGTGAAGTGAAGTGAAG 1740
Db 1681 AACTGGCAGAAATTTGTCAGATGCAAAATTTGAAGTGAAGTGAAGTGAAGTGAAG 1740
Oy 1741 AAATTAAGTAAAGACCACTGGAGTGTATTAATGAGATGATTAAGTGTAGTGTCT 1800
Db 1741 AAATTAAGTAAAGACCACTGGAGTGTATTAATGAGATGATTAAGTGTAGTGTCT 1800
Oy 1801 ACTTGGAAAGTGGGTGTATGATGCTACTGAAACCAATCAAGAACCCATTTAAGAA 1860
Db 1801 ACTTGGAAAGTGGGTGTATGATGCTACTGAAACCAATCAAGAACCCATTTAAGAA 1860
Oy 1861 CTCTTAGAGGCAAGAGAGCTATGTTGAGATCAGAAATCTTATGCGTGAATGGTGA 1920
Db 1861 CTCTTAGAGGCAAGAGAGCTATGTTGAGATCAGAAATCTTATGCGTGAATGGTGA 1920
Oy 1921 GCGGCTAGGCTCCATGAGAGCTGTAATCTCAACTCACTTTGGATCTCAATGAGT 1980
Db 1921 GCGGCTAGGCTCCATGAGAGCTGTAATCTCAACTCACTTTGGATCTCAATGAGT 1980
Oy 1981 GGTGAAGAGTCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2040
Db 1981 GGTGAAGAGTCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2040
Oy 2041 ATACCTTAGAGGAGTCCGCGACCAAACTGACCGAGGATGAGTTCGACATGTTTG 2100
Db 2041 ATACCTTAGAGGAGTCCGCGACCAAACTGACCGAGGATGAGTTCGACATGTTTG 2100
Oy 2101 GCTTGTGATGAGAGAGATCCACATGCGCTTTCCTGGAAGTGTATCCAGAAC 2160
Db 2101 GCTTGTGATGAGAGAGATCCACATGCGCTTTCCTGGAAGTGTATCCAGAAC 2160
Oy 2161 ACATGATTAATCTCAAGGCTTCAATCAATCACTGAGTAACCAATGTTTACGCT 2220
Db 2161 ACATGATTAATCTCAAGGCTTCAATCAATCACTGAGTAACCAATGTTTACGCT 2220
Oy 2221 CCATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2280
Db 2221 CCATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2280
Oy 2281 TGTCTAAATTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2340
Db 2281 TGTCTAAATTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2340
Oy 2341 ACTTGGGAGAAATTAATTAAGTGAATGATGATGATGATGATGATGATGATGATGAT 2396
Db 2341 ACTTGGGAGAAATTAATTAAGTGAATGATGATGATGATGATGATGATGATGATGAT 2396

```

RESULT 2
 PCT-US02-24048A-5
 ; Sequence 5, Application PC/TUS0224048A
 ; GENERAL INFORMATION:
 ; APPLICANT: E.I. du Pont de Nemours & Company

```

; TITLE OF INVENTION: Genes Involved in the Biosynthesis of Isopentenyl Diphosphate
; FILE OF INVENTION: brasiliensis latex
; FILE REFERENCE: CL1792 PCT
; CURRENT APPLICATION NUMBER: PCT/US02/24048A
; PRIOR APPLICATION DATE: 2002-09-10
; PRIOR FILING DATE: 2001-07-25
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO: 5
; LENGTH: 1509
; TYPE: DNA
; ORGANISM: Hevea brasiliensis
; PCT-US02-24048A-5

Query Match 35.1%; Score 840.2; DB 1; Length 1509;
Best Local Similarity 73.1%; Pred. No. 1.9e-207;
Matches 1108; Conservative 0; Mismatches 398; Indels 9; Gaps 2;

Oy 685 ATGGCTGTGTTGCTTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 744
Db 1 ATGGCTGTGTTGCTTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 60
Oy 745 GAGAACCCAAATGCAGGCGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 804
Db 61 GAGAACCCAAATGCAGGCGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 120
Oy 805 CCATTAACAGAGAGTCAAGCTGAAAGTGGGCTGGAATGGACAGATGCAATTA 864
Db 121 CCATTAACAGAGAGTCAAGCTGAAAGTGGGCTGGAATGGACAGATGCAATTA 180
Oy 865 ACATCCACAGCTCTCGAGAGAAAGCATGTAACTGCACTGATGATGATGATGATGATGATGATGAT 924
Db 181 ACATCCACAGCTCTCGAGAGAAAGCATGTAACTGCACTGATGATGATGATGATGATGATGATGAT 240
Oy 925 CAGTGTGTGCTGCAAGTATTAAGAAACCCCTTTGTAGAGATGCGATGATGATGATGATGATGATGAT 984
Db 241 CAGTGTGTGCTGCAAGTATTAAGAAACCCCTTTGTAGAGATGCGATGATGATGATGATGATGATGAT 300
Oy 985 ATAGCTGCTGCTATTTGGCAACCGAAGG---ACAAAGATCAATGGCAACACCTTA 1041
Db 301 ATAGCTGCTGCTATTTGGCAACCGAAGG---ACAAAGATCAATGGCTTAAACAGCTACTC 360
Oy 1042 TTGCAAGGCTGATGATTAACAAATTAAGGCTCAATGATGATGATGATGATGATGATGATGATGAT 1101
Db 361 TTGCAAGGCTGATGATTAACAAATTAAGGCTCAATGATGATGATGATGATGATGATGATGATGAT 420
Oy 1102 ATGAATGCGGCTGCTTCCATTTGACACGAGATGCGTGGTACCTTGCACCGTTTGA 1161
Db 421 ATGAATGCGGCTGCTTCCATTTGACACGAGATGCGTGGTACCTTGCACCGTTTGC 480
Oy 1162 TCAATCACTTAATGCTCGGAGTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1221
Db 481 TCAATCACTTAATGCTCGGAGTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
Oy 1222 ACTGGCTGATGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1281
Db 541 ACTGGCTGATGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
Oy 1282 CTGGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1341
Db 601 CTGGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 654
Oy 1342 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1401
Db 655 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 714
Oy 1402 GGGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1461
Db 715 GGGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 774
Oy 1462 GTCTGTCAATTTGCTGAGGCTGACAGTACAGGCTGCGCATTAATGAAGTATTTGGTACA 1521

```


Db 1015 GCAATTCAGACCTTGAACCAATTCATATTTTAAGCAAGCTCGCAGAAAGACATTGG 1074
Qy 1762 GATGTTTTCACAGCTATTAAGCTGTGTAGTGTCTTACTTTCGAAAAGTGGTGTTA 1821
Db 1075 GACGGTATTAATGTGTATAGACAGTGTGACAGCAAAAACTCAGAGAGGTGATGAG 1134
Qy 1822 CATGCTACTGAACCAATCAAGAACCATTTATTAAGAAGCTTTAGAGCAAGACAGCT 1881
Db 1135 CAGCAACTGAACCCAGCCGAGAACAGTGTGTTAAAGCATTTATTAAGTCAAGAAATGCC 1194
Qy 1882 ATGTTGAGAGTCAAGAAATCTTATGCGTCAGATGGGTGAGGGCGGTAGCGTTCCGATAGAG 1941
Db 1195 ATGCTTCAGATCAGAAATTAACATGCGCAGATGGGTGAGGGCTGCGAGTTCGATAGAG 1254
Qy 1942 CCTGAATTCACAACTTCTTGTGATCTTCAATGAGTGTCTGAGAGAGTCTTACTTCT 2001
Db 1255 CCTGAATTCACAGACTCGCTTGTGATCTTCAATGAGTGTGAGAGAGTCTTGTGCT 1314
Qy 2002 GGTGTTCTGAGCTGTGAGATTTGATGCCATTTTGCATTCATCTTGGGATTCGCGC 2061
Db 1315 GGAGTCTCTGAGCAGAGGTGGTGTGATGCTTCTGCTGTACTTATAGGGAGCTCTGGT 1374
Qy 2062 ACCAACTGACCCAGGAGTGTGAGTGTGACAAATGTTTGGCTTGTGAGAGAGAT 2121
Db 1375 ACCAATGTGCAAAAGCTTGTGAGTGTGACTCAATGTTTGGCTTGTGAGAGAGAC 1434
Qy 2122 CCACATGGCGTTTGGCTGAAAGTGTGATTCACAGAACCAATGATTAATCTTACAGCGTT 2181
Db 1435 CCTAATGTTGTTTGTGTAAGAGCGGCGATCCAGAACCAAGAAATCAACACAGCTGT 1494
Qy 2182 TCATCAATTCACCTT 2196
Db 1495 TTTGACGTTCAATTT 1509

RESULT 4
US-09-803-736-1155/C
Sequence 1155, Application US/09803736
GENERAL INFORMATION:
APPLICANT: Bush, David F.
APPLICANT: Levin, Irena M.
APPLICANT: Norris, Susan R.
APPLICANT: Rounsley, Steven D.
APPLICANT: Wiegand, Roger C.
TITLE OF INVENTION: Plant polymorphic Markers and Uses Thereof
FILE REFERENCE: 38-10115493D
CURRENT APPLICATION NUMBER: US/09/803,736
PRIOR FILING DATE: 2001-03-12
PRIOR APPLICATION NUMBER: US 09/534,859
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: Identified by Attorney Docket number 04983, 0206CPUS01 38-10
NUMBER OF SEQ ID NOS: 1582
SEQ ID NO 1155
LENGTH: 119420
TYPE: DNA
ORGANISM: Arabidopsis thaliana
US-09-803-736-1155

Query Match 31.7%; Score 759; DB 31; Length 119420;
Best Local Similarity 72.6%; Pred. No. 1,76-185;
Matches 1283; Conservative 0; Mismatches 0; Indels 484; Gaps 4;

Qy 1100 AGATAGATCGGCTGGCTTCCATTGACAGAAATCGCTGGGTACCTTGCACCGTTTG 1159
Db 29078 AGATAGATCGGCTGGCTTCCATTGACAGAAATCGCTGGGTACCTTGCACCGTTTG 29019
Qy 1160 CATCAATCAATTAATCTCGAGTCAAAATGTTAATTTCCAAAGCTGTAAGTAGCA 1219
Db 29018 CATCAATCAATTAATCTCGAGTCAAAATGTTAATTTCCAAAGCTGTAAGTAGCA 28959
Qy 1220 AAATGCTTAAAGTCTTCTGACAGCAATGACAAAGCTGTGGTTCACAGCTCTGTACAT 1279

Db 28958 AAATGCTTAAAGTCTTCTGACAGCAATGACAAAGCTGTGGTTCACAGCTCTGTACAT 28899
Qy 1280 ATCTTGAGTGTGTACTATCTATCTGATCAATGTAAGAAAGAAAGTTGGCTTGTGATC 1339
Db 28898 ATCTTGAGTGTGTACTATCTATCTGATCAATGTAAGAAAGAAAGTTGGCTTGTGATC 28839
Qy 1340 TAGATGTTATCCATATGATGATGACAAAGCTCATGTTGTTGCAAGGAAAGTGGCA 1399
Db 28838 TAGATGTTATCCATATGATGATGACAAAGCTCATGTTGTTGCAAGGAAAGTGGCA 28779
Qy 1400 GTGGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1459
Db 28778 GTGGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 28719
Qy 1460 AAGTCTTGTCAATTCCTC----- 1477
Db 28718 AAGTCTTGTCAATTCCTCAGTTCAATTAACAGAAAGTGTGTATATTAATTCCTCCTA 28659
Qy 1478 -----AGTTGCAATAGAGTGGCCATTAA 1504
Db 28658 AGAAAGGAATTTTACTACTCCTCCAAAGTTTGTAGGTTGCAATAGAGTGTGCTTAA 28599
Qy 1505 ATGAAGTATTTGATCAATTTTGAAGGAAATGAGCAATTAAGAACTGATGCTCTT 1564
Db 28598 ATGAAGTATTTGATCAATTTTGAAGGAAATGAGCAATTAAGAACTGATGCTCTT 28539
Qy 1565 TACCACTGATGATGATCTT----- 1584
Db 28538 TACCACTGATGATGATCTTGAACAAACAAATCAAGACTTAATTTCTAGTAAAGAAAC 28479
Qy 1585 ----- 1584
Db 28478 CTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 28419
Qy 1585 -----TTCCTTGAGAACTGGAAGTGTGATCCTCCACACATC 1625
Db 28418 AAGATGTTGCTTATTTAGTTCCTTGAGAACTGGAAGTGTGATCCTCCACACATC 28359
Qy 1626 AATGTAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1685
Db 28358 AATGTAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 28299
Qy 1686 GCAGAAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1745
Db 28298 GCAGAAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 28239
Qy 1746 AGCTAAAGCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1805
Db 28238 AGCTAAAGCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 28179
Qy 1806 TGAAA----- 1810
Db 28178 TGAAAAGGTATATTTCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 28119
Qy 1811 -----AGTGGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1835
Db 28118 CAGTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 28059
Qy 1836 AATCAAGCAAGCAATTAATTAAGAACTTTAGAGCAAGAAAGTATGTTGAGATGATGATGATGATGATGATGATGATGAT 1895
Db 28058 AATCAAGCAAGCAATTAATTAAGAACTTTAGAGCAAGAAAGTATGTTGAGATGATGATGATGATGATGATGATGATGAT 27999
Qy 1896 AATCTTATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1934
Db 27998 AATCTTATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 27939
Qy 1935 ----- 1934
Db 27938 TTGATAGTTGCTTCAATCTCAAGCAAGCAATTTTGTGATGATGATGATGATGATGATGATGATGATGAT 27879
Qy 1935 ----- 1934

Db 27878 AATAAATCTGCTGCTGATGATCTTAACAGCATATGATATACGGGTTCCACAAT 27819
QY 1935 -----G 1935
Db 27818 TCCGATTTAGAACAAATTGACTTCATCATTAATTGACCTCGAGGACCATTGAATTGACG 27759
QY 1936 ATAGAGCTGATCTCAAACTCACTTTTGATCTCTCAATGAGTGGGAGGATCTTA 1995
Db 27758 ATAGAGCTGATCTCAAACTCACTTTTGATCTCTCAATGAGTGGGAGGATCTTA 27699
QY 1996 CTTCCTGCTGCTGCTGAGTGTGATTTGATCCATATTGCAATCACTTTGAGGAGT 2055
Db 27698 CTTCCTGCTGCTGCTGAGTGTGATTTGATCCATATTGCAATCACTTTGAGGAGT 27639
QY 2056 TCCGGACCAAACTGACCCAGGACATGAGATTGCGACAATGTTTGCCCTGTTGGTGA 2115
Db 27638 TCCGGACCAAACTGACCCAGGACATGAGATTGCGACAATGTTTGCCCTGTTGGTGA 27579
QY 2116 GAATATCCACATGCGCTTGGCTGAGAAAGTGGTATCCAGAACACATGATTAATTCA 2175
Db 27578 GAATATCCACATGCGCTTGGCTGAGAAAGTGGTATCCAGAACACATGATTAATTCA 27519
QY 2176 GGCCTTTCATCAATTCACCTGATGATTAACAACATTTGTTAGTGTCCATATTAGGTG 2235
Db 27518 GGCCTTTCATCAATTCACCTGATGATTAACAACATTTGTTAGTGTCCATATTAGGTG 27459
QY 2236 GTCACCAAGTGGTGTGATGATGATGTTTGGATATGACATTTGGGTGTAATTTCTGG 2295
Db 27458 GTCACCAAGTGGTGTGATGATGATGTTTGGATATGACATTTGGGTGTAATTTCTGG 27399
QY 2296 TGTAGATTTTATACCATTGTAAGTCTTTTAACCTTTGGAACCTTTGGGGAATA 2355
Db 27398 TGTAGATTTTATACCATTGTAAGTCTTTTAACCTTTGGAACCTTTGGGGAATA 27339
QY 2356 AATAAGTGTATTCAAATCTTCTCA 2382
Db 27338 AATAAGTGTATTCAAATCTTCTCA 27312

RESULT 5
US-09-654-617-121976
; Sequence 121976, Application US/09654617
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Liu, Jindong
; TITLE OF INVENTION: Annotated Plant Genes
; FILE REFERENCE: 38-21(15097)D
; CURRENT APPLICATION NUMBER: US/09/654,617
; CURRENT FILING DATE: 2000-09-05
; NUMBER OF SEQ ID NOS: 463173
; SEQ ID NO 121976
; LENGTH: 576
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-654-617-121976

Query Match 23.8%; Score 569.6; DB 25; Length 576;
Best Local Similarity 99.3%; Pred. No. 3.9e-137;
Matches 572; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Db 181 GTTCTGCTGCTTCTCCGAGCTCCAGGCTAGTTGGTTTATTTTCAAGATTTTGC 240
QY 301 TTCTCTTTCATCGCGCAGCAGCAGCTGATTTCTGTCAAAACGTTAAGATTCGACTC 360
Db 241 TTCTCTTTCATCGCGCAGCAGCAGCTGATTTCTGTCAAAACGTTAAGATTCGACTC 300
QY 361 GAGGCTGACAGTAAGAGAGACAGCAGCTTGTGTGATGATGACGCGCAACGTGTGT 420
Db 301 GAGGCTGACAGTAAGAGAGACAGCAGCTTGTGTGATGATGACGCGCAACGTGTGT 360
QY 421 CGATCCGCTGCTGATCGAGCGAGTAATGATTCGATTCGATTCGATTCGATTCGATTCG 480
Db 361 CGATCCGCTGCTGATCGAGCGAGTAATGATTCGATTCGATTCGATTCGATTCGATTCG 420
QY 481 TTGAGATCTAATTCGAAATTCGATTAATATCATCTCCATCTCTTCTGGAAGTCCG 540
Db 421 TTGAGATCTAATTCGAAATTCGATTAATATCATCTCCATCTCTTCTGGAAGTCCG 480
QY 541 AATCCGATCTACACACACTACTGCTACCGCGGTATTTACTGCGCGGATTTCAATTA 600
Db 481 AATCCGATCTACACACACTACTGCTACCGCGGTATTTACTGCGCGGATTTCAATTA 540
QY 601 TCCGATCATTTCCGCGGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTC 636
Db 541 TCCGATCATTTCCGCGGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTC 576

RESULT 6
US-09-684-016-121976
; Sequence 121976, Application US/09684016
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Liu, Jindong
; TITLE OF INVENTION: Annotated Plant Genes
; FILE REFERENCE: 38-21(15097)D
; CURRENT APPLICATION NUMBER: US/09/684,016
; CURRENT FILING DATE: 2000-10-10
; PRIOR APPLICATION NUMBER: US 09/654,617
; NUMBER OF SEQ ID NOS: 2000-09-05
; SEQ ID NO 121976
; LENGTH: 576
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-684-016-121976

Query Match 23.8%; Score 569.6; DB 27; Length 576;
Best Local Similarity 99.3%; Pred. No. 3.9e-137;
Matches 572; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 61 AAGCAAACTGGGGAAGATGAGATGAGTGAAGAAACAAACCGTATTAACCTTGGGT 120
Db 1 AAGCAAACTGGGGAAGATGAGATGAGTGAAGAAACAAACCGTATTAACCTTGGGT 60
QY 121 TCAGAGGTGCGGAACGACGACCGCTTAACGAAATCTCAAAAGAAATTCGCGATCG 180
Db 61 TCAGAGGTGCGGAACGACGACCGCTTAACGAAATCTCAAAAGAAATTCGCGATCG 120
QY 181 GTTCTCTTTCATCGCGGACGACGAGTGTGCTGTAACGTTTACGATTCGACTC 240
Db 121 GTTCTCTTTCATCGCGGACGACGAGTGTGCTGTAACGTTTACGATTCGACTC 180
QY 241 GTTCTCTTTCATCGCGGACGACGAGTGTGCTGTAACGTTTACGATTCGACTC 300

QY 61 AAGCAAACTGGGGAAGATGAGATGAGTGAAGAAACAAACCGTATTAACCTTGGGT 120
Db 1 AAGCAAACTGGGGAAGATGAGATGAGTGAAGAAACAAACCGTATTAACCTTGGGT 60
QY 121 TCAGAGGTGCGGAACGACGACCGCTTAACGAAATCTCAAAAGAAATTCGCGATCG 180
Db 61 TCAGAGGTGCGGAACGACGACCGCTTAACGAAATCTCAAAAGAAATTCGCGATCG 120
QY 181 GTTCTCTTTCATCGCGGACGACGAGTGTGCTGTAACGTTTACGATTCGACTC 240
Db 121 GTTCTCTTTCATCGCGGACGACGAGTGTGCTGTAACGTTTACGATTCGACTC 180
QY 241 GTTCTCTTTCATCGCGGACGACGAGTGTGCTGTAACGTTTACGATTCGACTC 300
Db 241 GTTCTCTTTCATCGCGGACGACGAGTGTGCTGTAACGTTTACGATTCGACTC 300
QY 361 GAGGCTGACAGTAAGAGAGACAGCAGCTTGTGTGATGATGACGCGCAACGTGTGT 420

Db 301 GAGCGTCACAGTAAGAGAAGACAGCGATTGTGTAGATGACGCGCAACGTGTG 360
 Oy 421 CGATCCGCTGCTGATGACGAGGAATAGCTTTCATCGGTTTCGATCCAAATCGGAGT 480
 Db 361 CGATCCGCTGCTGATGACGAGGAATAGCTTTCATCGGTTTCGATCCAAATCGGAGT 420
 Oy 481 TTGAGGATCTAAATGGAATTTGCAATTAATCTCATCTCATCTCTCTGAGAGTCCG 540
 Db 421 TTGAGGATCTAAATGGAATTTGCAATTAATCTCATCTCATCTCTCTGAGAGTCCG 480
 Oy 541 AATCGATCTACACACTACTCGTACCGCGGCTCATTTACTGCCCGGATTTCAATTA 600
 Db 481 AATCGATCTACACACTACTCGTACCGCGGCTCATTTACTGCCCGGATTTCAATTA 540
 Oy 601 TCCGATCTTCCGCGGATTCGATCCAGACTGA 636
 Db 541 TCCGATCTTCCGCGGATTCGATCCAGACTGA 576

RESULT 7

US-10-219-999-9087
 : Sequence 9087, Application US/10219999
 : GENERAL INFORMATION:
 : APPLICANT: Cao, Yongwei
 : APPLICANT: Edgerton, Michael D
 : APPLICANT: Hinkle, Gregory J.
 : APPLICANT: Kovalic, David K.
 : APPLICANT: Liu, Jindong
 : APPLICANT: Stein, Joshua
 : TITLE OF INVENTION: CDNA SEQUENCES AND USES FOR PLANT IMPROVEMENT
 : FILE REFERENCE: 38-10(52726)C
 : CURRENT APPLICATION NUMBER: US/10/219, 999
 : PRIOR FILING DATE: 2002-08-15
 : PRIOR APPLICATION NUMBER: US 60/324,109
 : PRIOR FILING DATE: 2001-09-21
 : PRIOR APPLICATION NUMBER: US 60/312,544
 : PRIOR FILING DATE: 2001-08-15
 : NUMBER OF SEQ ID NOS: 63520
 : SEQ ID NO 9087
 : LENGTH: 1956
 : TYPE: DNA
 : ORGANISM: Zea mays
 : FEATURE:
 : NAME/KEY: CDS
 : LOCATION: (140)..(1675)
 : OTHER INFORMATION:
 : US-10-219-999-9087

Query Match 21.6%; Score 518.4; DB 42; Length 1956;
 Best Local Similarity 61.0%; Pred. No. 1.4e-123;
 Matches 941; Conservative 0; Mismatches 571; Indels 30; Gaps 5;

Oy 685 AMGGCTGTGTTGCTTCTGCTCTCGTGGGAAAGTTTGATGACAGCGGCTACTCTTACTC 744
 Db 140 ATGGAGAGTGGCTGCTGCGCGCGGGAAGGTGCTATCGCGGGAGATCTCTGTCTG 199
 Oy 745 GAGAAGCAATGACAGGCTTGTGTTGATCAATGACAGGCTTTAGCGATTTGAAG 804
 Db 200 GAGCGCCCAACGCGCGGCTGCTCTCAACACAGCGCGGCTTTAGCGCGCTCCGCC 259
 Oy 805 CCAATCAACGAAGAAGTCAAGCTGAAGTTGGGATGGAATGACAGATGCAATTA 864
 Db 260 CCCCCTCCCGACAGCTCTCCCGCCGACGCTGGGCTGGGCTGAGCGATGCAAAAGTG 319
 Oy 865 ACATCACACAGCTCTGAGAGAAAGCATATAAGTCACTCAATCATTTGACTCTT 924
 Db 320 ACTTCCCTCAGCTCTCAAGGAGGACACATACAAAGTGTCCCTAAGCAAGTCAGCGTG 379
 Oy 925 CAGTGTGTCTGCAAGATTCAGAAACCCCTTTAGAGCATGATACATATGCT 984
 Db 380 CAATTGACATCTGCGAGGGAATCTCAAAACCTTTTGTGGAGAGCCCATACATTTCTG 439
 Oy 985 ATAGCTGTGCTCATTTGGCAA-----CCGAGAAAGACAAAGATCATTTGCACAAACTC 1038

Db 440 GTAGCAGCTGCTAAAGCCGACCTATATGATAGAGAGAGAAAGATGATAGATTAAGTTG 499
 Oy 1039 TTATTCGAAGGCTCTGATATATACATATTTAGCTCCATATGATTTTACTCATATGGAAC 1098
 Db 500 CTGCTCAAGGCTCTAACAATTTACATTTCTTGATCAATGATTTTCTATTTGAAAG 559
 Oy 1099 CAGATAGATGGGCTGGGCTTCATTCAGACACAGATGCTGGGATACCTTGACCCGTTT 1158
 Db 560 CAGATTTGAAGCAGCGGCTCTCCCTCTTACTCCAGAGGCACTGCTTCAATGGCTCATTC 619
 Oy 1159 GCATCAATCACATTCATGCTGGAGTCAATGGTGTCT-----AATTCAG 1206
 Db 620 TCTTCAATCACATTTAATCAGAGGTGCTATGAGAACATGACGTGAGAGAAATGTA 679
 Oy 1207 CCTGAAGTAGCAAAACCTGGCTTAAAGTCTTCTGACGAATATACAAACAGCTGCTTGA 1266
 Db 680 CCTGAAGTAGCCAAACAGGACTTGGTCTATCAGCTGCCATGACCAATCACTGTTTGA 739
 Oy 1267 GCTGTGTACATTATCTTGAGAGTGGTGGCTGACATCTCATCTCAATGTA 1320
 Db 740 GCTTCTTCACTATCTTGATGCTGCTGATGCTCTCATGCTCCAGCAATCTTCTGTGAT 799
 Oy 1321 AAGTTGGCTGTTCTGATCTAGATGTTATCATATGATAGACAAACGTCATTTGCTT 1380
 Db 800 AATACACTAGACAGAACTTACTGTTGCTACCTGATGCCCCAAAGTGCACATTTGATA 859
 Oy 1381 GCACAAGGAAGTGGGAAGTGGTGTGATGCTGACCTGCTGCTGATGAGTACAGGT 1440
 Db 860 GCACAGGGGAAATTTGGAGCTGTTTGAATGTAGTCTGCTGTCTATGAGGATGACCG 919
 Oy 1441 TATGTTGCTGCTCTCCAGAACTGTTGTCATTTGCTCAAGTTGCAAGTTCAGCTGCA 1500
 Db 920 TATGTAAGTTTCTCCAGAAATACTCTGT---CTCAGGCTACAGGTGGGACTTTCTC 976
 Oy 1501 TTAATGAAGTTATTGTTCAATTTTGAAGGAAATTTGGACATTAAGAACTGAGTTC 1560
 Db 977 CCAGATGTAGTATCAAGAT---ATTGTTACACAAAGGTGGATCATGAGATTAACAGTTC 1033
 Oy 1561 TCTTTACACACAGCTAGTATGATTAATCTTTCTTGGAGAACCTGGAGAGTGGTCTCCACA 1620
 Db 1034 TCAATTAACCTCTCTGATGACCTTCTCTCTGAGGAACCTGGAGACCGGAGATCATCTACT 1093
 Oy 1621 CCATCAATGTTAGTGTGCTAGTAAGAAAGTGGCAATATGCTGATCCAGAGAGGACAGAA 1680
 Db 1094 CCATCAATGTTGATCTGTGTAAGAGGTGGAGAGTGTGACCCCTGAGAAATGCAAAAGT 1153
 Oy 1681 AACTGGCAGATTTGTGATGCAATTTAGAACTGGAACTTAAGCTTAACGATCTGAGC 1740
 Db 1154 ACATGGAGTAAACTGTGCTTCCCAATTCAGCGCTGGAGAACCACTGGCAATCTTAA 1213
 Oy 1741 AATTTAGCTAAAGACCACTGGAGTGTATCTACAGAGTATTAAGTCTGTGATGCTT 1800
 Db 1214 GGACTTTCTGAAATATCCGAGAGGATATGATGCTGTAGTCTATCTGTGCCATCTC 1273
 Oy 1801 ACTTGTGAAAGTGGGTGTTTACATGCTACTGCAACCAATCAAGAACCAATTAATAAGAA 1860
 Db 1274 ACTTATGAGAAATGGACAGAGGTGCTACTAACAACATCAAGAAATTAATTAATACA 1333
 Oy 1861 CTCTTAGAGGCAAGAGACTATGTTGAGATCAGAAATCTTATGCTGATAGTGGTGA 1920
 Db 1334 TTGTTGGCTGCAAGGATGCTTGTGCTGAGAAAGGCTTCAATGAGGAGAGATGGCATA 1393
 Oy 1921 GGGGTAGGCTTCGATAGAGCTGAATCTCAACCTTTGGATTCCTACATGAGT 1980
 Db 1394 GCAGCGGGTTCGATAGGACAGATTCACAAACAGGCTACTTGACGCCACTATGAAAT 1453
 Oy 1981 GCTGAAGAGTTCTACTGCTGCTGCTGCTGCTGAGAGCTGTGATTTGATCCATATTTGCA 2040
 Db 1454 ATGAGAGGGTGTTTACTTCTGAGATTCGGAGGCTGTGGCTTGATGATCAATATTTCA 1513
 Oy 2041 ATCACTTAGGGGATTCGGGACCAAACTGACCCAGGCAATGAGTGTGCAACAATGTTTTC 2100


```

US-09-649-165A-2752/C
; Sequence 2752, Application US/09649165A
; GENERAL INFORMATION:
; APPLICANT: Guegler, Karl J.
; APPLICANT: Doyle, Martin
; APPLICANT: Momiyama, Monika
; APPLICANT: Wang, Xinhao
; TITLE OF INVENTION: DETECTING EXPRESSION OF ARABIDOPSIS THALIANA GENES
; FILE REFERENCE: PA-0019 US
; CURRENT APPLICATION NUMBER: US/09/649,165A
; CURRENT FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: 60/150,517
; PRIOR FILING DATE: 1999-08-24
; NUMBER OF SEQ ID NOS: 7753
; SOFTWARE: PERL Program
; SEQ ID NO 2752
; LENGTH: 475
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No: 701516116
US-09-649-165A-2752

Query Match
Best Local Similarity 99.8%; Score 473.4; DB 25; Length 475;
Pred. No. 4.2e-112;
Matches 474; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1874 GAGAAGCTATGTTGAGATAGAAATTTTATGCTGATGAGTGGGCTGAGCGCTTC 1933
DB 475 GAGAAGCTATGTTGAGATAGAAATTTTATGCTGATGAGTGGGCTGAGCGCTTC 416
QY 1934 CGATGAGCTGATGATCAACATCACTTTGATTTAGCAATGAGTGGGCTGAGCGCTTC 1993
DB 415 CGATGAGCTGATGATCAACATCACTTTGATTTAGCAATGAGTGGGCTGAGCGCTTC 356
QY 1994 TACTTGTGTTGTTCTGAGAGTGGTGGATTTGATGCAATTTGCAATGACTTTAGCGG 2053
DB 355 TACTTGTGTTGTTCTGAGAGTGGTGGATTTGATGCAATTTGCAATGACTTTAGCGG 296
QY 2054 ATTCCGGCACCACCACTGAGCCAGGATGAGTGGCAATGTTTGGCTTGTGTGTA 2113
DB 295 ATTCCGGCACCACCACTGAGCCAGGATGAGTGGCAATGTTTGGCTTGTGTGTA 236
QY 2114 GAGAAGATCCACATGGCGTTGGCTAGAAAGTGGTATCCAGCAACACATGTTTACTT 2173
DB 235 GAGAAGATCCACATGGCGTTGGCTAGAAAGTGGTATCCAGCAACACATGTTTACTT 176
QY 2174 CAGGCGTTTCATCAATTCACCTTGAGTAACACATGTTTCACTGATTTAGT 2233
DB 175 CAGGCGTTTCATCAATTCACCTTGAGTAACACATGTTTCACTGATTTAGT 116
QY 2234 GCGTCACCAAGTTGCGTTGAGTATCTGTTTGCATATGACTGGTGGCTAAATTTCTT 2293
DB 115 GCGTCACCAAGTTGCGTTGAGTATCTGTTTGCATATGACTGGTGGCTAAATTTCTT 56
QY 2294 GGTGTAGACATTTTATACCATTTGTAAGCTTTTAACTCTTGAAGAACTGGCG 2348
DB 55 GGTGTAGACATTTTATACCATTTGTAAGCTTTTAACTCTTGAAGAACTGGCG 1

RESULT 10
US-09-708-427-83406
; Sequence 83406, Application US/09708427
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1243P
; CURRENT APPLICATION NUMBER: US/09/708,427
; CURRENT FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 85364
; SOFTWARE: PatentIn version 3.1

```

```

; SEQ ID NO 83406
; LENGTH: 1889
; TYPE: DNA
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 1..1889
; OTHER INFORMATION: any n = a, g, c, t, unknown, or other
; NAME/KEY: misc.feature
; LOCATION: 1..1889
; OTHER INFORMATION: Ceres Seq. ID 1962603
US-09-708-427-83406

Query Match
Best Local Similarity 59.8%; Score 463.6; DB 28; Length 1889;
Pred. No. 2.7e-109;
Matches 922; Conservative 0; Mismatches 554; Indels 66; Gaps 6;

QY 685 ATGGCTGTTGTTGCTCTCTGGAAGTTTGTGATGAGTGGAGCTTACTTGTACTC 744
DB 111 ATGGAAGTGGTGGCTGGCGCCGGGAGAGTGCTCATCGGGAGGATACCTGCTGCTG 170
QY 745 GAGAGGCAATATGAGGCTGTGTGTGATCAATGACAGCTTTTACGGCATTTGTAAG 804
DB 171 GAGGCGCCCAAGCGGCTGATGCTCAGCACACCGCCGCTTACGCGCTGCTCCG 230
QY 805 CCAATCAACGAGAAGTCAAGCTGAAAGTTGGCATGAAATGAGACATGCAATTA 864
DB 231 CCCCTCCGCGACAGACCTCCCGCGAGCTTGGGCTGGGCTGGAGATGTCAAAAGTG 290
QY 865 ACATCACCACAGCTCTCGAGAGAAAGCATATTAAGTGTATCAATGATTTGACTCTT 924
DB 291 ACTTCCCTCAGCTCTCAGGAGAGCCACATACAAAGCTGTCCATACGAAGTGGCGCTG 350
QY 925 CAGTCTGTGTGCAATGATTTCAAGAACCCCTTTGTAGAGCATGCGATACATAGCT 984
DB 351 CAATTTGACATCTGGAGGAGATTCACAAACCTTTTGTGGAGGAGGCGCATCAATTTCTG 410
QY 985 ATAGCTGCTGCTCAATTTGGCA-----CCGAGAGAGCAAGATCATTTGCACAACTC 1038
DB 411 GTACAGCTGTAAAGGACCATTTATGATAGAGAGAGAAAGATGTAGTATAGTGTG 470
QY 1039 TTATTTGCAAGCTCTTGTATATACATATTTAGGCTTCATGACTTTTACTATTCGGAAC 1098
DB 471 CTGCTGCAAGTGTGAACATATTAATTTCTGATGATGATGATTTCTATATAGAAAG 530
QY 1099 CAGATTAATGCGTGGGCTTCATTTGACACCAATGCGTGGGTACCTTGCAACGTTT 1158
DB 531 CAGATTTGAAGCAGCGGCTTCCCTTACTTCCAGAGGACACGCTTTTCAATGCCATTC 590
QY 1159 GCATCAATCAATCAATGCTGGGAGTCAAA-----TGGTGTATTTCCAAAG 1206
DB 591 TCTTCAATCAATTTAACTCAGAGGTTGCTATATGGAACATGACTGGAGAGAAATGTAA 650
QY 1207 CCTGAAGTAGCAAAACTGCGTTAGGTTCTTCTGACAGCAATGACAAAGCTGTGTGCA 1266
DB 651 CCTGAAGTAGCAAAACTGCGTTAGGTTGCTATGACCTGACATGACACATCAGTTGTGCA 710
QY 1267 GCTCTGTATCTATTTCTTGAGTGTGACCTATCTGATTCATTTAA-----GAAGGA 1320
DB 711 GCTCTTCTTCAATCTTGTGGTGTGCTTACCTTCATGCTCAGAGCAATCTTCTGGGAT 770
QY 1321 AAGTTGGCTGTTCTGATCTGATGATATATCATTGATAGCAAAAGCTCATGTTCTT 1380
DB 771 AATTAACATGAGCAGAACTTGACTTGTGCTTCTATTTGCCCAAGTGCACATTTGATA 830
QY 1381 GCACAGGAGAGGTCGAGAGTGGTTGATGTCAGCTGTGCTGTATGAGAGTCAAGCT 1440
DB 831 GCACAGGAGAGGTCGAGAGTGGTTGATGTCAGCTGTGCTGTATGAGAGTCAAGCT 890
QY 1441 TATGTTGGCTTCTCTCAGAGATGTTGATGATTTCTAGAGTGTGAGTAACAGCTGCCA 1500
DB 891 TATGTAAGGTTTCTCCAGAAATATCTCTGT--CTCAGGCTACAGGTGGAGCTTCTCTC 947

```


1661 AACGGCAGATTTCAGATGCAATTTAGACGAGAACTAACCTAAACGATCTGAC 1740
1123 ACATGAGTAAAGTCCCTTCCATTCAGCGCTGAGAAACCACTGGCAATCTTAA 1182
1741 AATAGTAAAGCACTGGATTTATCTAGACATCATTAATCTTGTGTGTGCT 1800
1183 GACTTTCTGAAATACCGAGAG----- 1207
1801 ACTTGTAAAGTGGGTTCATCATCTACGACCAATCAACGAGACCTATTAAGA 1860
1208 -CATGATGCTGTGAGAGAGTGGCTACTATACCAATCAAGATTAATTTAAAT 1266
1861 CTCTTAGAGGCAAGAACTATGTTGAGATCAGAAATTTCTTATGCTGAG 1920
1267 TTGTTGCTGAGAGAGATGCTTGCCTGAGATTAAGCTTACATCCAGAGATG 1326
1921 GCGGTACGTTCCATGAGAGCTTAATCTCAACTTTTGGATTTCTAATGAGT 1980
1327 GCAGCGGTGTCCAGATGAGCAATTCACAAACGCTTACCTTACCTATGAA 1386
1981 GCTGAAGAGTTCTACTGCTGTGCTGCTGAGAGTGGATTTGATGACATATT 2040
1387 ATGAGAGGTGTTTACTTCTGATGATTCGAGAGCTGGTGGCTTGATGACAT 1446
2041 ATCACTTTAGGGGATTCGCGACCAACTGACCGAGATGAGTTCGACAAATG 2100
1447 GTGACTTTGGGGATGCAAGTAATCTGTAGCAAAATGCTTGGAGCTGCTGG 1506
2101 GCCTTTGTTGAGAGAGATCCACATGCGCTTGGCTAGAAAGTGTGATCCAG 2160
1507 CCACCTTCGCTGTCGAGAGACTGCGGTGCTCATTTGGAAGATGCTGACCA 1566
2161 ACATGATTTACTTTCAGCGCTTTCATCAATTCACCTTGAGTAA 2202
1567 AGGAGATGATCGGACGCTGATGCTCATTCAGATTATTTGA 1608

RESULT 12
US-09-638-111-871/c

; Sequence 871, Application US/09638111
; GENERAL INFORMATION:
; APPLICANT: G.Tlach, J.Tm
; TITLE OF INVENTION: Expressed Sequences of Arabidopsis
; FILE REFERENCE: 2008US
; CURRENT APPLICATION NUMBER: US/09/638,111
; PRIORITY FILING DATE: 2000-08-11
; PRIOR APPLICATION NUMBER: US 60/148,783
; NUMBER OF SEQ ID NOS: 900
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 871
; LENGTH: 443
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-638-111-871

Query Match 18.5%; Score 443; DB 24; Length 443;

Best Local Similarity 100.0%; Pred. No. 3.4e-104; Indels 0; Gaps 0;

Matches 443; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1896 AATTCCTATGCGTCAGATGAGGCGGCTAGCGTTCGATAGAGCCGATCTCAAC 1955
443 AATTCCTATGCGTCAGATGAGGCGGCTAGCGTTCGATAGAGCCGATCTCAAC 384
1956 TCAACTTTGATTTACAAATGAGTGTGAGAGAGTTCTACTGCTGTGCTTGGAGC 2015
383 TCAACTTTGATTTACAAATGAGTGTGAGAGAGTTCTACTGCTGTGCTTGGAGC 324
2016 TGGTGGATTTGATGCAATATTTGCAATCACTTTAGGGGATTCGCGACCAAACTGACCA 2075
323 TGGTGGATTTGATGCAATATTTGCAATCACTTTAGGGGATTCGCGACCAAACTGACCA 264

2076 GGCATGAGAGTTCGACCAATGTTTGGCTTGTGTGAGAGAGATCATAGCGCTT 2135
263 GCATGAGAGTTCGACCAATGTTTGGCTTGTGTGAGAGAGATCATAGCGCTT 204
2136 CCTAGAAATGATGATCCAGCAACATGTTTACTTATGAGCGTTTCAATTCACCT 2195
203 CCTAGAAATGATGATCCAGCAACATGTTTACTTATGAGCGTTTCAATTCACCT 144
2196 TGAGTAAACAAATGTTTACTTATGAGCGTTCACCAATGAGTGGTGGT 2255
143 TGAGTAAACAAATGTTTACTTATGAGCGTTCACCAATGAGTGGTGGT 84
2256 ATACTGTTTGGATATGAGAGTGGGCTTAAATTTCTTGTGTGATCAATTTTAA 2315
83 ATACTGTTTGGATATGAGAGTGGGCTTAAATTTCTTGTGTGATCAATTTTAA 24
2316 TTGTAAGTCTTAACTTTGCA 2338
23 TTGTAAGTCTTAACTTTGCA 1

RESULT 13
US-09-770-444-851/c

; Sequence 851, Application US/09770444
; GENERAL INFORMATION:
; APPLICANT: Gorlach, Jörn
; APPLICANT: An, Yong-Oiang
; APPLICANT: Hamilton, Carol M.
; APPLICANT: Price, Jennifer L.
; APPLICANT: Raines, Tracy M.
; APPLICANT: Yu, Yang
; APPLICANT: Rameaka, Joshua G.
; APPLICANT: Page, Amy
; APPLICANT: Matthew, Abraham V.
; APPLICANT: Ledford, Brooke L.
; APPLICANT: Woessner, Jeffrey P.
; APPLICANT: Haas, William David
; APPLICANT: Garcia, Carlos A.
; APPLICANT: Krickler, Maja
; APPLICANT: Slader, Ted
; APPLICANT: Davis, Keith R.
; APPLICANT: Allen, Keith
; APPLICANT: Hoffman, Neil
; APPLICANT: Hurban, Patrick
; TITLE OF INVENTION: Expressed Sequences of Arabidopsis
; FILE REFERENCE: 2027 (PARA-016PRV)
; CURRENT APPLICATION NUMBER: US/09/770,444
; PRIORITY FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 60/178,502
; NUMBER OF SEQ ID NOS: 999
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 851
; LENGTH: 443
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-770-444-851

Query Match 18.5%; Score 443; DB 30; Length 443;

Best Local Similarity 100.0%; Pred. No. 3.4e-104; Indels 0; Gaps 0;

Matches 443; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1896 AATTCCTATGCGTCAGATGAGGCGGCTAGCGTTCGATAGAGCCGATCTCAAC 1955
443 AATTCCTATGCGTCAGATGAGGCGGCTAGCGTTCGATAGAGCCGATCTCAAC 384
1956 TCAACTTTGATTTACAAATGAGTGTGAGAGAGTTCTACTGCTGTGCTTGGAGC 2015
383 TCAACTTTGATTTACAAATGAGTGTGAGAGAGTTCTACTGCTGTGCTTGGAGC 324
2016 TGGTGGATTTGATGCAATATTTGCAATCACTTTAGGGGATTCGCGACCAAACTGACCA 2075

Db 323 TGGTGATTTGATGCGAATTTGCAATACCTTTAGGGGATTCGGACCAAACTGACCCA 264
QY 2076 GGCATGAGATTGCGACAAATGTTTGGCCTTGTGGTGAGAGAAGATCCACATGGCGTTTG 2135
Db 263 GGCATGAGATTGCGACAAATGTTTGGCCTTGTGGTGAGAGAAGATCCACATGGCGTTTG 204
QY 2136 CCTGAAGAAGTGTGATCCAGACACATGTAATTAATTCAGAGCGTTTATCAATTCACCT 2195
Db 203 CCTGAAGAAGTGTGATCCAGACACATGTAATTAATTCAGAGCGTTTATCAATTCACCT 144
QY 2196 TGAGTAACAACATTTGTTAGTGTGATGATTAATTTAGTGTGCGTACACAGTTCGGTTGAGT 2255
Db 143 TGAGTAACAACATTTGTTAGTGTGATGATTAATTTAGTGTGCGTACACAGTTCGGTTGAGT 84
QY 2256 ATACTGTTTGCATATAGACATTTGGTGTAAATTTCTTGGTGAAGCATTTTATATACCA 2315
Db 83 ATACTGTTTGCATATAGACATTTGGTGTAAATTTCTTGGTGAAGCATTTTATATACCA 24
QY 2316 TTGTAAGTCTTTAACTCTTGGA 2338
Db 23 TTGTAAGTCTTTAACTCTTGGA 1

RESULT 14

US-09-654-617-192867
: Sequence 192867, Application US/09654617
: GENERAL INFORMATION:
: APPLICANT: Kovalic, David K.
: APPLICANT: Liu, Jindong
: TITLE OF INVENTION: Annotated Plant Genes
: FILE REFERENCE: 38-21(15097)D
: CURRENT APPLICATION NUMBER: US/09/654,617
: CURRENT FILING DATE: 2000-09-05
: NUMBER OF SEQ ID NOS: 463173
: SEQ ID NO 192867
: LENGTH: 728
: TYPE: DNA
: ORGANISM: Gossypium hirsutum
: OTHER INFORMATION: unsure at all n locations
US-09-654-617-192867

Query Match 17.8%; Score 425.4; DB 25; Length 728;
Best Local Similarity 76.1%; Pred. No. 1.6e-99;
Matches 554; Conservative 0; Mismatches 162; Indels 12; Gaps 2;

QY 784 CGGTTTACGGATTTGAGCCATCAACAGAGTCAAGCCTGGAAGTTGGGCAATGG 843
Db 1 CGTTTTCAGCTATTTGTAAGCAATCATGAGCTATCAAGCCTGGAAGCTGGGCAATGG 60
QY 844 AAATGACAGATGTCAAATTAACATCACACAGCTCTCGAGAGAAAGCATGTATTAACCTG 903
Db 61 TCTTGAGACGATGTCAAATTAACATCACACAGCTCTCGAGAGAAAGCATGTATTAACCTG 120
QY 904 TCACGATTCATTTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 963
Db 121 TCTGGAACATTTAATCACTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
QY 964 GAGCATGCGATACAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020
Db 181 GAAATGCTATTAATCACTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240
QY 1021 GAATCATTCGACAACTCTTATTCGAAAGCTTGTATATTAACAAATTAAGCTTCACATGAC 1080
Db 241 GAGCATGCGATACAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
QY 1081 TTTTACATCATTCGGAACCAATAGAAATCGGCTGGCTTCATGACACCGAATCGCTG 1140
Db 301 TTTTACATCATTCGGAACCAATAGAAATCGGCTGGCTTCATGACACCGAATCGCTG 360
QY 1141 GGTACCCCTGGACCGTTTGCATCATTCATGCTGCGGAGTCAAAATGGTGTAT 1200
Db 361 GGTACCCCTGGACCGTTTGCATCATTCATGCTGCGGAGTCAAAATGGTGTAT 420

QY 1201 TCCAACTGTAAGTAGCAAAAACCTGCTTACGTTCTCTGACGACATGACAAACCTGTG 1260
Db 421 TGCAAACTGTAAGTAGCAAAAACCTGCTTACGTTCTCTGACGACATGACAAACCTGTG 480
QY 1261 GTTGACGCTGCTTACATTAATTCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1315
Db 481 GTTGACGCTGCTTACATTAATTCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
QY 1316 ---AAGAAAGTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1371
Db 541 CACCAAGAAATTAAGAAATTCACAGATCTCGATATGTCATGATGATGATGATGATGATGATG 600
QY 1372 CATGCTCTGCAACAAGGAGAGTGGGCTTATGTCAGCTGCTGCTGCTGCTGCTGCTGCTG 1431
Db 601 CATGCTCTGCAACAAGGAGAGTGGGCTTATGTCAGCTGCTGCTGCTGCTGCTGCTGCTG 660
QY 1432 AGTCAGCTTATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1491
Db 661 AGTCAGCTTATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
QY 1492 GGTCTGCC 1499
Db 721 GGTCTGCC 728

RESULT 15

US-09-684-016-192867
: Sequence 192867, Application US/09684016
: GENERAL INFORMATION:
: APPLICANT: Kovalic, David K.
: APPLICANT: Liu, Jindong
: TITLE OF INVENTION: Annotated Plant Genes
: FILE REFERENCE: 38-21(15097)D
: CURRENT APPLICATION NUMBER: US/09/684,016
: CURRENT FILING DATE: 2000-10-10
: PRIOR APPLICATION NUMBER: US 09/654,617
: PRIOR FILING DATE: 2000-09-05
: NUMBER OF SEQ ID NOS: 463173
: SEQ ID NO 192867
: LENGTH: 728
: TYPE: DNA
: ORGANISM: Gossypium hirsutum
: FEATURE:
: NAME/KEY: unsure
: LOCATION: (1)..(728)
: OTHER INFORMATION: unsure at all n locations
US-09-684-016-192867

Query Match 17.8%; Score 425.4; DB 27; Length 728;
Best Local Similarity 76.1%; Pred. No. 1.6e-99;
Matches 554; Conservative 0; Mismatches 162; Indels 12; Gaps 2;

QY 784 CGGTTTACGGATTTGAGCCATCAACAGAGTCAAGCCTGGAAGTTGGGCAATGG 843
Db 1 CGTTTTCAGCTATTTGTAAGCAATCATGAGCTATCAAGCCTGGAAGCTGGGCAATGG 60
QY 844 AAATGACAGATGTCAAATTAACATCACACAGCTCTCGAGAGAAAGCATGTATTAACCTG 903
Db 61 TCTTGAGACGATGTCAAATTAACATCACACAGCTCTCGAGAGAAAGCATGTATTAACCTG 120
QY 904 TCACGATTCATTTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 963
Db 121 TCTGGAACATTTAATCACTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
QY 964 GAGCATGCGATACAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020
Db 181 GAAATGCTATTAATCACTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240
QY 1021 GAATCATTCGACAACTCTTATTCGAAAGCTTGTATATTAACAAATTAAGCTTCACATGAC 1080
Db 241 GAGCATGCGATACAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
QY 1081 TTTTACATCATTCGGAACCAATAGAAATCGGCTGGCTTCATGACACCGAATCGCTG 1140

```
Db 301 TTTACTCATACAGAAATCAGATAGAAAGCACTTGGTCTTCGGTTGACACCTGAACCATTTG 360
QY 1141 GGTACCCCTTGACACCGTTTGCATCAATCAATTCATGCTGCGAGTCAAAATGCTCAAT 1200
Db 361 GCTACTCTACACCGTTTACATCAATTCATCAATTCCTGAGGATCAAAATGAGCAAAAT 420
QY 1201 TCCAGACCTGAAGTAGCAAAAGCTGCTTGTCTCTGAGCAAAATGACAAACAGCTGTG 1260
Db 421 TGCAAACCTGAAGTTGCAAAAGCTGATGCTGATCATCTGACAAATGACAACTGCTGA 480
QY 1261 GTTCAGCTCTGTACATTAATCTGAGTGGTGGTACCTATCTGATCATGTAAAG----- 1315
Db 481 GTTGTGCTTACTTCATTAATCTGCTGTGTGTTAACCTTTCACCTCTTTCGACATCA 540
QY 1316 ----AAGGAAGTTGGCTGTCTGATCTAGATGTTATCCATATGATAGCACAAAGCTCT 1371
Db 541 CACCAGAAATTAAGAAATCCACAGATCTGATATGTGATATGATAGCTCAAAAGTCC 600
QY 1372 CATGTCTTGCACAAAGGAGGTGAGTGGTGTGATGTCAGCTGCTGTATGGA 1431
Db 601 CACTGTATTTGCCAAGAGTAAAGTTGGCAGTGGCTTGAATGTCAGTTCTGTATGGG 660
QY 1432 AGTCAGCGTATGTTGCTCTCTCCAGAGTCTTGTCAATTTGCTCAGGTTGCAGTAACA 1491
Db 661 AGTCAGCGTATGTTGCTCTCTCCAGAGTCTTGTCTGCTGCTCAGGCTGCANTGAAA 720
QY 1492 GGTCTGCC 1499
Db 721 GGGATGCC 728
```

Search completed: May 3, 2003, 21:11:50
Job time : 3833 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 3, 2003, 19:40:12 ; Search time 190 Seconds
(without alignments)
14900.864 Million cell updates/sec

Title: US-09-988-863A-1

Perfect score: 2396
Sequence: 1 gtcgaccacgcgtccggc.....tttccaaaaa.....2396

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 746064 seqs, 590810554 residues

Total number of hits satisfying chosen parameters: 1492128

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications_NA:*

1: /cgn2-6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2-6/ptodata/1/pubpna/PCR_NEW_PUB.seq:*
3: /cgn2-6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2-6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2-6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2-6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
7: /cgn2-6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2-6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2-6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
10: /cgn2-6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgn2-6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
12: /cgn2-6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
13: /cgn2-6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
14: /cgn2-6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	2396	100.0	2396	10	US-09-988-863A-1
2	443	18.5	443	10	US-09-770-444-851
3	425.4	17.8	728	10	US-09-988-863A-4
4	299.2	12.5	611	10	US-09-988-863A-3
5	229.2	9.6	757	10	US-09-909-745-21
6	150.4	6.3	539	10	US-09-909-745-19
7	135.2	5.6	571	10	US-09-988-863A-5
8	123.2	5.1	2004	10	US-09-887-576-273
9	107.2	4.5	249	10	US-09-909-745-15
10	107.2	4.5	249	10	US-09-909-745-17
11	87.4	3.6	1816	9	US-09-938-842A-4887
12	76.6	3.2	2000	10	US-09-887-576-204
13	57.6	2.4	1356	9	US-09-918-740-61
14	57.6	2.4	7681	9	US-09-918-740-61
15	57.6	2.4	7693	9	US-09-918-740-58
16	57.6	2.4	7695	9	US-09-918-740-59
17	57.6	2.4	8224	9	US-09-918-740-62
18	57.6	2.4	8235	9	US-09-918-740-60
19	57.6	2.4	8400	9	US-09-918-740-64

20	57.6	2.4	13917	9	US-09-918-740-72	Sequence 72, Appl
21	57.6	2.4	14623	9	US-09-918-740-74	Sequence 74, Appl
22	57.6	2.4	14623	9	US-09-918-740-76	Sequence 76, Appl
23	49.6	2.1	2000	9	US-09-938-842A-4507	Sequence 4507, Ap
24	40.8	1.7	486	10	US-09-878-574-4606	Sequence 4606, Ap
25	39	1.6	1523	9	US-09-938-842A-4708	Sequence 4708, Ap
26	38	1.6	27681	9	US-10-091-504-1997	Sequence 1997, Ap
27	38	1.6	27681	9	US-10-091-504-1998	Sequence 1998, Ap
28	38	1.6	27681	10	US-09-764-869-1997	Sequence 1997, Ap
29	38	1.6	27681	10	US-09-764-869-1998	Sequence 1998, Ap
30	37.4	1.6	2080	9	US-09-991-936-1867	Sequence 1867, Ap
31	37.4	1.6	2080	9	US-09-991-936-1869	Sequence 1869, Ap
32	37.2	1.6	368	10	US-09-834-975-47	Sequence 47, Appl
33	37.2	1.6	32190	9	US-10-079-854-201	Sequence 201, App
34	37.2	1.6	32190	10	US-09-764-878-201	Sequence 201, App
35	37.2	1.6	32193	9	US-10-079-854-200	Sequence 200, App
36	37.2	1.6	32193	10	US-09-764-878-200	Sequence 200, App
37	37.2	1.6	32249	9	US-10-079-854-202	Sequence 202, App
38	37.2	1.6	32249	10	US-09-764-878-202	Sequence 202, App
39	37	1.5	65042	9	US-10-229-124-3	Sequence 3, Appl1
40	36.8	1.5	380	10	US-09-960-352-9335	Sequence 9335, Ap
41	36.8	1.5	386	10	US-09-829-481-5	Sequence 5, Appl1
42	36.6	1.5	941	9	US-10-123-155-464	Sequence 464, App
43	36.2	1.5	55155	10	US-09-735-933-3	Sequence 3, Appl1
44	35.8	1.5	443	10	US-09-770-444-842	Sequence 842, App
45	35.4	1.5	401	10	US-09-960-352-10213	Sequence 10213, A

ALIGNMENTS

RESULT 1
US-09-988-863A-1
Sequence 1, Application US/09988863A
Patent No. US20020123427A1
GENERAL INFORMATION:
APPLICANT: Bayer AG
TITLE OF INVENTION: Plant phosphomevalonate kinases
FILE REFERENCE: Le A 35 018
CURRENT APPLICATION NUMBER: US/09/988, 863A
CURRENT FILING DATE: 2001-11-21
NUMBER OF SEQ ID NOS: 5
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 2396
TYPE: DNA
ORGANISM: Arabidopsis thaliana
FEATURE:
NAME/KEY: CDS
LOCATION: (685)..(2199)
US-09-988-863A-1

Query Match 100.0%; Score 2396; DB 10; Length 2396;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2396; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTGACCCACGCGTCCGGCGGCGACCTCTCTCTCTCTTACGACACCAATGATAG 60
DB 1 GTGACCCACGCGTCCGGCGGCGACCTCTCTCTCTCTTACGACACCAATGATAG 60
QY 61 AAGCAAACTGGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 120
DB 61 AAGCAAACTGGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 120
QY 121 TCAGAGGTGCGGACCGACCGACCGACCGACCGACCGACCGACCGACCGACCGAC 180
DB 121 TCAGAGGTGCGGACCGACCGACCGACCGACCGACCGACCGACCGACCGACCGAC 180
QY 181 GTTCT 240
DB 181 GTTCT 240
QY 241 GTTCT 300
DB 241 GTTCT 300

|||||
Db 241 GTTCTGGTTCTTCGAACTCCAGGCCAGTTGGTTTATTTTTCACAGATTTGTC 300
OY 301 TTCTCTTTCATGCGGCGACGACGAGTCTGATCTCTGCAAAAGTTAAAGCATCGACTC 360
Db 301 TTCTCTTTCATGCGGCGACGACGAGTCTGATCTCTGCAAAAGTTAAAGCATCGACTC 360
OY 361 GAGCGTCACAGTAAGAGAAAGACAGCGATGTTGTGTAGATGACGGCGAGGTGTGT 420
Db 361 GAGCGTCACAGTAAGAGAAAGACAGCGATGTTGTGTAGATGACGGCGAGGTGTGT 420
OY 421 CGATCCGTCCTGATGACGAGGAGATAGTTTGCATCCGCTTTCATCCAAATGCGAGCT 480
Db 421 CGATCCGTCCTGATGACGAGGAGATAGTTTGCATCCGCTTTCATCCAAATGCGAGCT 480
OY 481 TTGAGATCTTAATCGAAATTTGATTAATCTCATCTCCAAATCTCTTCTGAAGAGTCG 540
Db 481 TTGAGATCTTAATCGAAATTTGATTAATCTCATCTCCAAATCTCTTCTGAAGAGTCG 540
OY 541 AATCCGATCTACACACACTACTGTCACCGCGGTCAATTTACTGCGCGCATTTCAATTA 600
Db 541 AATCCGATCTACACACACTACTGTCACCGCGGTCAATTTACTGCGCGCATTTCAATTA 600
OY 601 TCCGATATTTCCGCGCATATCCAAATGCGACAGTGAATGAGGTGTTTGTATGACG 660
Db 601 TCCGATATTTCCGCGCATATCCAAATGCGACAGTGAATGAGGTGTTTGTATGACG 660
OY 661 GATATCTTTGTCACCTCTTTCGAAATGAGTGTGCTTTCGCTTCCGCGAAAGTTTG 720
Db 661 GATATCTTTGTCACCTCTTTCGAAATGAGTGTGCTTTCGCTTCCGCGAAAGTTTG 720
OY 721 ATGACGTGAGGCTACTTGTACTCGAAGCCAAATGCAAGGCTTGTGTGATGACAAAT 780
Db 721 ATGACGTGAGGCTACTTGTACTCGAAGCCAAATGCAAGGCTTGTGTGATGACAAAT 780
OY 781 GCACGTTTACGCGATTTGCAAGCCAAATGCAAGGCTTGTGTGATGACAAAT 840
Db 781 GCACGTTTACGCGATTTGCAAGCCAAATGCAAGGCTTGTGTGATGACAAAT 840
OY 841 TGGAAATGACAGATGTCAAATTAACATCAACACAGCTCTGAGAGAAAGCATTTAA 900
Db 841 TGGAAATGACAGATGTCAAATTAACATCAACACAGCTCTGAGAGAAAGCATTTAA 900
OY 901 CTGTACATGAATCTTTGACTCTTCACTGTGTGTGCAAGTATTCAGAAACCCCTTT 960
Db 901 CTGTACATGAATCTTTGACTCTTCACTGTGTGTGCAAGTATTCAGAAACCCCTTT 960
OY 961 GTAGAGCATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
Db 961 GTAGAGCATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
OY 1021 GAATCATTTGCAAACTCTTATTTGCAAGGCTTGTATTAACAAATTAAGGCTCCAAATGAC 1080
Db 1021 GAATCATTTGCAAACTCTTATTTGCAAGGCTTGTATTAACAAATTAAGGCTCCAAATGAC 1080
OY 1081 TTTTATCTATATGCGAAGCAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
Db 1081 TTTTATCTATATGCGAAGCAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
OY 1141 GGTACCCCTTGACCGTTTGCATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 1200
Db 1141 GGTACCCCTTGACCGTTTGCATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 1200
OY 1201 TCCAAAGCTGAAGTGAAGAAAGTGGCTTCTTCTGAGAGATGACAAAGCTGTG 1260
Db 1201 TCCAAAGCTGAAGTGAAGAAAGTGGCTTCTTCTGAGAGATGACAAAGCTGTG 1260
OY 1261 GTTACAGCTCTGTTACATTAATCTTGAAGTGTGATGATGATGATGATGATGATGATGATGAT 1320
Db 1261 GTTACAGCTCTGTTACATTAATCTTGAAGTGTGATGATGATGATGATGATGATGATGATGAT 1320
OY 1321 AAGTTGGCTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1380
Db 1321 AAGTTGGCTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1380

Db 1381 AAGTTGGCTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1380
OY 1381 GCACAGGAAAGTTCGGAAGTGGCTTGTATGATGATGATGATGATGATGATGATGATGATGATGAT 1440
Db 1381 GCACAGGAAAGTTCGGAAGTGGCTTGTATGATGATGATGATGATGATGATGATGATGATGATGAT 1440
OY 1441 TATGTTGCTTCTTCGAGAAATGCTTGTATGATGATGATGATGATGATGATGATGATGATGATGAT 1500
Db 1441 TATGTTGCTTCTTCGAGAAATGCTTGTATGATGATGATGATGATGATGATGATGATGATGATGAT 1500
OY 1501 TTAATGAGTATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1560
Db 1501 TTAATGAGTATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1560
OY 1561 TCTTTACACACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1620
Db 1561 TCTTTACACACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1620
OY 1621 CCATCAATGTTAGTGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1680
Db 1621 CCATCAATGTTAGTGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1680
OY 1681 AACTGGAGATTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1740
Db 1681 AACTGGAGATTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1740
OY 1741 AATTTAGCTAAAGACACTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1800
Db 1741 AATTTAGCTAAAGACACTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1800
OY 1801 ACTTCTGAAAAGTGGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1860
Db 1801 ACTTCTGAAAAGTGGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1860
OY 1861 CTCTTAGAGGCAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1920
Db 1861 CTCTTAGAGGCAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1920
OY 1921 GCGGCTAGCGCTTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1980
Db 1921 GCGGCTAGCGCTTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1980
OY 1981 GCTAAGAGATTTCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2040
Db 1981 GCTAAGAGATTTCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2040
OY 2041 ATCACTTTAGGGGATTCGCGACCAAACTGACCCAGGATGATGATGATGATGATGATGATGAT 2100
Db 2041 ATCACTTTAGGGGATTCGCGACCAAACTGACCCAGGATGATGATGATGATGATGATGATGAT 2100
OY 2101 GCGTTGTTGTTGAGAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2160
Db 2101 GCGTTGTTGTTGAGAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2160
OY 2161 ACATGATATTTACTGAGGCTTTCATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 2220
Db 2161 ACATGATATTTACTGAGGCTTTCATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 2220
OY 2221 CCAATTTATAGGTCGTCACCAAACTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 2280
Db 2221 CCAATTTATAGGTCGTCACCAAACTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 2280
OY 2281 TGTAAATTTCTTGTGTAGAGATTTTATCAATGATGATGATGATGATGATGATGATGATGATGAT 2340
Db 2281 TGTAAATTTCTTGTGTAGAGATTTTATCAATGATGATGATGATGATGATGATGATGATGATGAT 2340
OY 2341 ACTTGGGGAATAAATAAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2396
Db 2341 ACTTGGGGAATAAATAAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2396

RESULT 2
US-09-770-444-851/c

Sequence 851, Application US/09770444
Patent No. US20020023280A1
GENERAL INFORMATION:
APPLICANT: Gorlach, Jorn
APPLICANT: An, Yong-Qiang
APPLICANT: Hamilton, Carol M.
APPLICANT: Price, Jennifer L.
APPLICANT: Raines, Tracy M.
APPLICANT: Yu, Yang
APPLICANT: Rameaka, Joshua G.
APPLICANT: Page, Amy
APPLICANT: Mathew, Abraham V.
APPLICANT: Ledford, Brooke L.
APPLICANT: Moessner, Jeffrey P.
APPLICANT: Haas, William David
APPLICANT: Garcia, Carlos A.
APPLICANT: Krieker, Maja
APPLICANT: Slader, Ted
APPLICANT: Davis, Keith R.
APPLICANT: Allen, Keith
APPLICANT: Hoffman, Neil
APPLICANT: Hurban, Patrick
TITLE OF INVENTION: Expressed Sequences of Arabidopsis
TITLE OF INVENTION: thailana
FILE REFERENCE: 2027 (PARA-016PRV)
CURRENT APPLICATION NUMBER: US/09/770,444
PRIORITY FILING DATE: 2001-01-26
PRIORITY APPLICATION NUMBER: 60/178,502
NUMBER OF SEQ ID NOS: 999
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 851
LENGTH: 443
TYPE: DNA
ORGANISM: Arabidopsis thailana
US-09-770-444-851

Query Match 18.5%; Score 443; DB 10; Length 443;
Best Local Similarity 100.0%; Pred. No. 2,8e-121;
Matches 443; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1896 AATTCCTATGCGTCAGATGGGTGAGCGGCTAGCGCTTCGATAGAGCCGTGAATCTCAAC 1955
DB 443 AATTCCTATGCGTCAGATGGGTGAGCGGCTAGCGCTTCGATAGAGCCGTGAATCTCAAC 384
QY 1956 TCACTTTTGGATTTCAACAATGAGTGTGAGAGAGTTCTACTTGGTGTCTTGGAGC 2015
DB 383 TCACTTTTGGATTTCAACAATGAGTGTGAGAGAGTTCTACTTGGTGTCTTGGAGC 324
QY 2016 TGGTGATTTGATGCGCATATTTGCAATCACTTAGGGGATTCGGGACCAACCTGACCA 2075
DB 323 TGGTGATTTGATGCGCATATTTGCAATCACTTAGGGGATTCGGGACCAACCTGACCA 264
QY 2076 GGCATGAGGTCGACAAATGTTTGGCTGTGTTGGTGAAGAATCCATGAGGCTTGG 2135
DB 263 GGCATGAGGTCGACAAATGTTTGGCTGTGTTGGTGAAGAATCCATGAGGCTTGG 204
QY 2136 CCTAAGAGTGTGATCCACGACCAACATGTTACTTACAGCGCTTATCAATTCACCT 2195
DB 203 CCTAAGAGTGTGATCCACGACCAACATGTTACTTACAGCGCTTATCAATTCACCT 144
QY 2196 TGAGTAACAACATGTTTCACTGTCAATTTATAGTGTGCGCACCAAGTTGGTTGAGT 2255
DB 143 TGAGTAACAACATGTTTCACTGTCAATTTATAGTGTGCGCACCAAGTTGGTTGAGT 84
QY 2256 ATACTGTTTGCATATAGACTGGGTGCTAAATTTCTGGTGAAGCATTTTATACCA 2315
DB 83 ATACTGTTTGCATATAGACTGGGTGCTAAATTTCTGGTGAAGCATTTTATACCA 24
QY 2316 TTGTAGGCTTTACTCTTGA 2338
DB 23 TTGTAGGCTTTACTCTTGA 1

RESULT 3
US-09-988-863a-4
Sequence 4, Application US/09988863A
Patent No. US20020123427A1
GENERAL INFORMATION:
APPLICANT: Bayer AG
TITLE OF INVENTION: Plant phosphomevalonate kinases
FILE REFERENCE: Le A 35 018
CURRENT APPLICATION NUMBER: US/09/988,863A
PRIORITY FILING DATE: 2001-11-21
NUMBER OF SEQ ID NOS: 5
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 4
LENGTH: 728
TYPE: DNA
ORGANISM: Gossypium hirsutum
FEATURE:
NAME/KEY: unsure
LOCATION: 715
OTHER INFORMATION: n can be any nucleotide
US-09-988-863a-4

Query Match 17.8%; Score 425.4; DB 10; Length 728;
Best Local Similarity 76.1%; Pred. No. 6.8e-116;
Matches 554; Conservative 0; Mismatches 162; Indels 12; Gaps 2;

QY 784 CGGTTTACGCGATTTGGAAGCCATCAACGAAAGTAAACCTGGAAGTGGCGATG 843
DB 1 CGTTTTACGCGATTTGGAAGCCATCAACGAAAGTAAACCTGGAAGTGGCGATG 60
QY 844 AATGAGACAGATGTCAAATTAATCAACCAAGCTCTGAGAGAAAGCATGTATTAAC 903
DB 61 TCTTGACCGAGATGTCAGATTAACATCTCTGCTTCCAGGAAAGCATGTATTAAC 120
QY 904 TCATGTAATCATTTGACTTCACTTCACTGCTGTGTCGAAGTGAAGAAACCCCTTGA 963
DB 121 TCTGGAAGCATTTTAACACTTCACTGCTGTGTCGAAGTGAAGAAACCCCTTGA 180
QY 964 GAGCATGCGATGACATGCTATGCTGCTGCTCATTTGGCA---ACGAGAAAGCAAA 1020
DB 181 GAAATGCTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 240
QY 1021 GAATGATTCGACAACTCTTATGCAAGGCTTGATATTAATTAATTAATTAATTA 1080
DB 241 GAGGCTTTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 300
QY 1081 TTTTACTATTCGGAACCAAGATAGATGAGCTGGGCTTCCATTGACACCAATGCTG 1140
DB 301 TTTTACTATTCGGAACCAAGATAGATGAGCTGGGCTTCCATTGACACCAATGCTG 360
QY 1141 GGTACCTTGGACCGCTTTCATCATCATCATCATCATCATCATCATCATCATCAT 1200
DB 361 GGTACCTTGGACCGCTTTCATCATCATCATCATCATCATCATCATCATCATCAT 420
QY 1201 TCCAGGCTGAGTGAAGAAACCTGCTAGGCTTTCGAGCAATGACACAGCTG 1260
DB 421 TCCAGGCTGAGTGAAGAAACCTGCTAGGCTTTCGAGCAATGACACAGCTG 480
QY 1261 GTTGAGCTCTGTTCAATATCTTGGAGTGTGACCTATCTGATCATGATGAAG 1315
DB 481 GTTGAGCTCTGTTCAATATCTTGGAGTGTGACCTATCTGATCATGATGAAG 540
QY 1316 ----AAGGAATTTGGCTGCTCTGATCTAGATGTTATTCATATGATACCAAGCT 1371
DB 541 CACCAAGAAATTAAGAAATTCACAGATTCGATATTTGCAATGATACCTAAAGTGC 600
QY 1372 CATGCTCTGCAAGGAGAGTGGAGAGTGGTGTGATGATGAGTGGCTCTATGA 1431
DB 601 CACTGATTTGCCAAGGATTAAGTGGAGTGGCTGTGATGATGATGATGATGATG 660
QY 1432 AGTCAGCGTTATGTTGCTCTCTCAGAGAGCTTGTGCTATTTGCTCAGGTTGCA 1491
DB 1432 AGTCAGCGTTATGTTGCTCTCTCAGAGAGCTTGTGCTATTTGCTCAGGTTGCA 1491

Db 661 AGTAGCGTTATGTTCTTTTACACCAAAAGTCTTCTGCTGCTGACGCTGCANTGAA 720
QY 1492 GGTCTGCC 1499
Db 721 GGATGCC 728

RESULT 4
US-09-988-863a-3
; Sequence 3, Application US/09988863A
; Patent No. US20020123427A1
; GENERAL INFORMATION:
; APPLICANT: Bayer AG
; TITLE OF INVENTION: Plant phosphomevalonate kinases
; FILE REFERENCE: Le A 35 018
; CURRENT APPLICATION NUMBER: US/09/988, 863A
; CURRENT FILING DATE: 2001-11-21
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 3
; LENGTH: 611
; TYPE: DNA
; ORGANISM: Medicago truncatula
US-09-988-863a-3

Query Match 12.5%; Score 299.2; DB 10; Length 611;
Best Local Similarity 71.7%; Pred. No. 2.2e-78;
Matches 407; Conservative 0; Mismatches 158; Indels 3; Gaps 1;

QY 691 GTTGTGCTGCTGCTGCTGGAAGTTTGTGACTGAGACCTGCTGCTGCTGAGAG 750
Db 44 GTTGTGCTGCTGCTGCTGGAAGTTTGTGACTGAGACCTGCTGCTGAGAG 103
QY 751 CCAATGCGAGGCTTGTGTGAGTACAATGACGCTTTTACGGATTGAGACCAATC 810
Db 104 CCTAATGCTGAGCTTGTGTGAGTACAATGACGCTTTTACGGATTGAGACCAATC 163
QY 811 AACCAAGAGTCAAGCTGGAAGTTGGCATGGAATGGACAGATGTCAAATTACATCA 870
Db 164 TATCTCTCAAACTAAACGATCTTGGCTTGGCTGAGATGTCAGATTAAACATCT 223
QY 871 CCACAGCTGCGAGAAAGATGATTAATGCTGACTGATTAATGCTGCTGCTGAGTCT 930
Db 224 CCTAATGCTGAGCTTGTGTGAGTACAATGACGCTTTTACGGATTGAGACCAATC 283
QY 931 GTGTGCTGAGTGTCAAGAAACCTTGTAGAGCATGGATACAGATGCTATAGCT 990
Db 284 GTTCTCAAGTGAAGCAAGAACCTTGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 343
QY 991 GCTGCTGATTTGGCAAC---CGAGAGGACAAAGATATGTCACAACTCTTATGCA 1047
Db 344 GCCGCTATGCAACAGCTGACCAAGATTAAGAAAGCTTGTGCAACAACTCTTATG 403
QY 1048 GGTGTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1107
Db 404 GGTGTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 463
QY 1108 TGGCTGAGGCTTGCATTAAGCAACAGATGCTGAGTACCTTGCACCTTGTGCAATC 1167
Db 464 AGACAGGAGCTCCCTTGCATTAAGCAACAGATGCTGAGTACCTTGTGCAATC 523
QY 1168 ACATTCAGATGCTGCGAGTCAATGCTGCTATTAATTCAGAGCTGAGAGAAAGTGGC 1227
Db 524 TCTTTCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 583
QY 1228 TTAGGTTCTGCTGAGCAATGACAAAG 1255
Db 584 TTGGGCTCATCTGACAGATGACAAAG 611

RESULT 5
US-09-909-745-21
; Sequence 21, Application US/09909745

; Patent No. US20020119546A1
; GENERAL INFORMATION:
; APPLICANT: Falco, S. Carl
; APPLICANT: Farnold, Omolayo O.
; TITLE OF INVENTION: Squalene Synthase Enzymes
; FILE REFERENCE: BB1112 US CIP
; CURRENT APPLICATION NUMBER: US/09/909, 745
; CURRENT FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: 60/107,241
; PRIOR FILING DATE: 1998-11-05
; PRIOR APPLICATION NUMBER: 60/107,241
; PRIOR FILING DATE: 1998-11-05
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 21
; LENGTH: 757
; TYPE: DNA
; ORGANISM: Oryza sativa
US-09-909-745-21

Query Match 9.6%; Score 229.2; DB 10; Length 757;
Best Local Similarity 61.3%; Pred. No. 1.8e-57;
Matches 369; Conservative 0; Mismatches 233; Indels 0; Gaps 0;

QY 1601 GAAGTGTGATCTCTCCACACATCAATGATGCTGAGTAAAGAGTGCATATCTCTG 1660
Db 9 GTATCGGAGGATCTCTCCACACATCAATGATGCTGAGTAAAGAGTGCATATCTCTG 68
QY 1661 ATCCAGAGAGGACGAGAAACCTGCGAGATTTGTGATGATGATGATGATGATGATG 1720
Db 69 ACCCTCAGAAATTCAGAGAGATGAGTAAATTTGGGATTTGATTTACGCTTGA 128
QY 1721 CTAAGCTAAAGATCTGAGCAATTAATTAAGTAAAGCAACGATGATGATGATGATG 1780
Db 129 ACCAAGTGAAGATCTAAACAACTGCTGAGTAAAGTAAAGTAAAGTAAAGTAAAGT 188
QY 1781 TTAAGTCTGAGTGTGCTTCTCTGAAAGTGGGTTTCACTGATGATGATGATGATG 1840
Db 189 TACGATCTGAGTGTGCTTCTCTGAAAGTGGGTTTCACTGATGATGATGATGATG 248
QY 1841 AGAAGCATTATTAAGAACTCTTAAAGCAAGAGATGATGATGATGATGATGATGATG 1900
Db 249 AAGAAGTATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 308
QY 1901 TTAAGTCTGAGTGTGCTTCTCTGAAAGTGGGTTTCACTGATGATGATGATGATG 1960
Db 309 ATATGCGAGAGATGAGGATGATGATGATGATGATGATGATGATGATGATGATG 368
QY 1961 TTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2020
Db 369 TTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 428
QY 2021 GATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2080
Db 429 GCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 488
QY 2081 GGAATGCGACAAATGTTTGGCTTGTGTGAGAGAAAGATGATGATGATGATGATG 2140
Db 489 GGAATGCGAGTGTGCTTCTCTGAAAGTGGGTTTCACTGATGATGATGATGATG 548
QY 2141 AAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2200
Db 549 AAGCTGTGAGCCCAAGAGAGAGGATGATGATGATGATGATGATGATGATGATG 608
QY 2201 AA 2202
Db 609 GA 610

RESULT 6
US-09-909-745-19
; Sequence 19, Application US/09909745
; Patent No. US20020119546A1

```

GENERAL INFORMATION:
APPLICANT: Falco, S. Carl
TITLE OF INVENTION: Squalene Synthase Enzymes
FILE REFERENCE: B1112 US CIP
CURRENT APPLICATION NUMBER: US/09/909,745
PRIORITY FILING DATE: 2001-07-20
PRIORITY APPLICATION NUMBER: 60/107,241
PRIORITY FILING DATE: 1998-11-05
PRIORITY APPLICATION NUMBER: 60/107,241
PRIORITY FILING DATE: 1998-11-05
NUMBER OF SEQ. ID NOS: 26
SOFTWARE: Microsoft Office 97
SEQ ID NO 19
LENGTH: 539
TYPE: DNA
ORGANISM: Oryza sativa
FEATURE:
NAME/KEY: unsure
LOCATION: (287)
NAME/KEY: unsure
LOCATION: (440)
NAME/KEY: unsure
LOCATION: (454)
NAME/KEY: unsure
LOCATION: (460)
NAME/KEY: unsure
LOCATION: (465)
NAME/KEY: unsure
LOCATION: (486)
NAME/KEY: unsure
LOCATION: (488)
NAME/KEY: unsure
LOCATION: (495)
NAME/KEY: unsure
LOCATION: (530)
NAME/KEY: unsure
LOCATION: (533)
US-09-909-745-19

```

```

DB          421  GGCTT  425

| | | |
| | | |

RESULT 7
US-09-988-863A-5
; Sequence 5, Application US/09988863A
; Patent No. US2002013427A1
; GENERAL INFORMATION:
; APPLICANT: Bayer AG
; TITLE OF INVENTION: Plant phosphomevalonate kinases
; FILE REFERENCE: Le A 35 018
; CURRENT APPLICATION NUMBER: US/09/988,863A
; CURRENT FILING DATE: 2001-11-21
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 571
; TYPE: DNA
; ORGANISM: Pinus radiata
; US-09-988-863A-5

```

OY	1721	CTAAGCTTAACAGACTGTGAGCAAAATTAGCTAAAGACCACCTGGGATGTTTATCTACAGCTGA	1780
Db	122	ACCAACTGAGGAACTTAACAACAACTTGCTGGAAGATCACGGGGAACCCATATGAATCTGTTT	181
OY	1781	TTAAGTCCTTAGTGTGCCTTACTCTTGA AAAAGTGCGTGTTCATGCTACTGACCAATCA	1840
Db	182	TACGATCCTGTATGTGCTCTCAAGTCCAGTAATGTGACAGAGGTGGCTACCATTCACAACTAC	241
OY	1841	ACGAAGCCATTATTAAGAAGCTCTTAGAGGCCAAGAGAAGCTATGTTGAGATCAGAAATTC	1900
Db	242	AAGAATAATTGTTATGATCATCTACTGCGCCGAAGAGATCTTCCNTGAAAATAAGGCTTC	301
OY	1901	TTATG-CGTCGATGTGGGTGAGGGCGCTAGCGCTTCGATGAGCCCTGAATCCTAAACTCAA	1959
Db	302	ATATGCCAAGAGATGGGATATGACAGCTGGTGTCCAAATTGAGCCGCAAGAACACAAACTCAA	361
OY	1960	CTTTTGAGATTCATACAGTAGTCTCAAGGAGTCTACTTGCTGGAGTCTCGAGACTGGT	2019
Db	362	CTTCTCGATGCCACATATGAATATGAGGGGTGTTCTACTACATACGG-ATTTCCTGGGGCGGCT	420
OY	2020	GGATT 2024	

RESULT 8
US-09-887-576-273/c
Sequence 273, Application US/09887576
Patent No. US20020144047A1
GENERAL INFORMATION:
APPLICANT: Budworth, P.
APPLICANT: Brown, D.
APPLICANT: Chang, H.
APPLICANT: Zhu, T.
APPLICANT: Han, B.
APPLICANT: Wang, X.
APPLICANT: Cooper, Bret
TITLE OF INVENTION: Promoters for regulation of plant expression
FILE REFERENCE: 1360.001us1
CURRENT APPLICATION NUMBER: US/09/887,576
CURRENT FILING DATE: 2001-06-25
PRIOR APPLICATION NUMBER: US 60/213,848
PRIOR FILING DATE: 2000-06-23
PRIOR APPLICATION NUMBER: US 60/214,087
PRIOR FILING DATE: 2000-06-23
PRIOR APPLICATION NUMBER: US 60/258,692
PRIOR FILING DATE: 2000-12-29

```

: NUMBER OF SEQ ID NOS: 875
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 273
: LENGTH: 2004
: TYPE: DNA
: ORGANISM: Arabidopsis thaliana
US-03-887-576-273

```

Query Match	5.1%;	Score 123.2;	DB 10;	Length 2004;
Best Local Similarity	-71.28;	Pred. No. 1.2e-25;		
Matches 195;	Conservative	0;	Mismatches 68;	Indels 11;
				Gaps 2

OY	271	TAGTGGTATTATTTACAGAGATTTTGCTCTTTATATGGGACACAGACGTCGA	330
Db	512	TAGTITTTTTTGGGCATTCACGCTGTGGTCTCTCGACGGCGGACG-TTTTTCAGGC	455
OY	331	GTTCCTGTCAAAAGCTTAACGATCCGACTCGAGCGTCGACAGTAAGAGAAAGACAGCG	390
Db	453	GATTGTATCTCTCGATATCGATCTGTCCTCATTTAGTCAAAAGGAAACAAACAGCGCA	394
OY	391	ATTGATGTC-----TAGATGAGCGGCAAGCGTGTGCATGCCCTCGATCGACGG	444
Db	333	ATTGTGTGGGCATTCGCTCTGCATCAACGGCGATGTGTGTGCATGCCGTCAATCGACGG	334
OY	441	AGAAATACGTTTTGATTCGCGGTTTCATCCAAATCGAGAGTTTGAGGATCTTAATCGGAAA	500
Db	333	CGATAGTGTGTGCATGTCGATCATTCACAAATCGAGAGTTTAAGGTTCCAAATCGGAGA	274
OY	501	TTGCATTAATCATCATCCCAATCTCTTCRAG	534
Db	273	ATTATATGATCATCTCCACATCTCTCTTAAG	240

RESULT 9
 US-09-909-745-15/c
 : Sequence 15, Application US/09909745
 : Patent No. US20020119546A1
 :
 : GENERAL INFORMATION:
 :
 : APPLICANT: Falco, S. Carl
 : APPLICANT: Famodu, Omolayo O.
 : TITLE OF INVENTION: Squelene Synthesis Enzymes
 : FILE REFERENCE: B0112 US CIP
 :
 : CURRENT APPLICATION NUMBER: US/09/909, 745
 :
 : PRIOR FILING DATE: 2001-07-20
 : PRIOR APPLICATION NUMBER: 60/107,241
 : PRIOR FILING DATE: 1998-11-05
 : PRIOR APPLICATION NUMBER: 60/107,241
 : PRIOR FILING DATE: 1998-11-05
 :
 : NUMBER OF SEQ ID NOS: 26
 :
 : SOFTWARE: Microsoft Office 97
 :
 : SEQ ID NO 15
 : : LENGTH: 249
 : : TYPE: DNA
 : : ORGANISM: Zea mays
 :
 : US-09-909-745-15

Query Match	4.5%	Score 107.2	DB 10	Length 249
Best Local Similarity	64.5%	Pred. No. 1.7e-21		
Matches 160	Conservative 0	Mismatches 88	Indels 0	Gaps 0

QY	1789	TGTTGTTGGCTTACTTCTTGAAAAGGTGGGTTCATGCTACAGCAACCAATCAAGAGGCC	1844
Db	249	TGTAGTGTCTTCACATATTTGGAAAGTGGGCAAGAGTGGCTATCAACCACATCAAAATAATTA	190
QY	1849	ATTATTAAGAACTCTTAGAGGCCAAGAAAGCTATGTAGAGATCAGAAATTCTTATGGCT	1908
Db	189	ATTATTAAGATCATCTTTGGCTTCGAAGGATGCTTGGCTTGATATGAAGGCTTACATGGCA	130
QY	1909	CAGATGGGTGAGCGGCTAGCGTTACGCGATAGAACCTGAAATCTCAAACTCAACTTTTGAT	1968
Db	129	GAGATGGGCATAGACAGCTGGTGTTCCAATTAGACCAGATATCAAAACAGCGCTACTAGAT	70
QY	1969	TCCTCAATAGATGCTCAAGCAATTCTACTTGGTGTCTTCCGAGACCTGGTGATTTGAT	2028

Db	QY	2029	2036
69	GCACATATGAAATGTSAGSGCTTTACTGGCTGGAGTTCCGSSAGCTGGTGGCTTCGAT	10	
9	GCAGTGT	2	

```

RESULT 10
US-09-909-745-17
Sequence 17, Application US/09909745
Patent No. US20020119546A1
GENERAL INFORMATION:
APPLICANT: Falco, S. Carl
APPLICANT: Famedu, Omolayo O.
TITLE OF INVENTION: Squalene Synthe
FILE REFERENCE: Bb1112 US CIP
CURRENT APPLICATION NUMBER: US/09/9
PRIOR FILING DATE: 2001-07-20
PRIOR APPLICATION NUMBER: 60/107,24
PRIOR FILING DATE: 1998-11-05
PRIOR APPLICATION NUMBER: 60/107,24
PRIOR FILING DATE: 1998-11-05
NUMBER OF SEQ ID NOS: 26
SOFTWARE: Microsoft Office 97
SEQ ID NO 17
LENGTH: 249
TYPE: DNA
ORGANISM: zea mays
US-09-909-745-17

```

Query Match	4.58;	Score 107.2;	DB 10;	Length 249;
Best Local Similarity	64.58;	Pred. No. 1.7e-21;		
Matches 160;	Conservative 0;	Mismatches 88;	Indels 0;	Gaps 0

OY	1789	TGTAGTGGCTTACTCTGAAAAGTGSGGTTCACATGACTGAACCAATCACCAGACC	1848
Dd	1	TGTAGTGGCTCACATATGSGAAGTGGCCAGAGTGGCTACTAACCAACATCACAATA	60
OY	1849	ATTATTAAAGACTCTTAGAGGCCAGAGAAGCTANGTTGAGATCAGATTCTATGGCT	1908
Dd	61	ATTATTAAAGATCATCTGTGTGGCTGCAMAGGATGCTTGCTTGACATTAAGCGCTTACAT	120
OY	1909	CAGATGGGTAGGGGGCTACCGTCCGATAGAGACCTGANTCTCAACATCACTTTGGAT	1968
Dd	121	GAGATGGGCATACGACGTGGTGTCCAATTGAGCCGATTCACMAACGGCTACTAAT	180
OY	1969	TCATCAATGAGTGGCTGAAGAGTTCTACTCTGCTGTGTTCTCTGAGCTGGTGGATTGAT	2028
Dd	181	GCCACTCTGGAATATGAGGGGTGTTTAACTGGCTGGAGTTCCGGGAGCTGGTGGCTTCAT	240
OY	2029	GCCATATT 2036	
Dd	241	GCACTGTT 248	

```

RESULT 11
US-09-938--842A-4887
: Sequence 4887, Application US/09938842A
: Patent No. US20020160378A1
: GENERAL INFORMATION:
: APPLICANT: Harper, Jeff
: APPLICANT: Kireps, Joel
: APPLICANT: Wang, Xun
: APPLICANT: Zhu, Tong
: TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
: TITLE OF INVENTION: SAME, AND METHODS OF USE
: FILE REFERENCE: SCRIPI300-3
: CURRENT APPLICATION NUMBER: US/09/938, 842A
: CURRENT FILING DATE: 2001-08-24
: PRIOR APPLICATION NUMBER: US 60/227,866
: PRIOR FILING DATE: 2000-08-24
: PRIOR APPLICATION NUMBER: US 60/264,647

```


Best Local Similarity 5.6%; Pred. No. 0.0038;
Matches 23; Conservative 213; Mismatches 177; Indels 0; Gaps 0;

```
OY 1494 TCTGCATTAATGAGTATTGTACATTTGAGGGAATGGGACATAGAGAC 1553
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1470 TATCAAGTACTTAAGAGATAGAGATTTGTCRRRRRRRRRRRRRRRRRR
OY 1554 TGAGTCTCTTACACCATGATGAACTTTCTTCTTGAGAACCTGGAATGGTGC 1613
      : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1410 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1351
OY 1614 CTCACACCATCATGTAGTAGTCAGTAAGAGTGGCAATGCTGTATCCAGAGAGGC 1673
      : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1350 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1291
OY 1674 ACAGAAACAGCAGATTTGTCAGTGAATTTAGAACCTGGAATAGCTAAAGCA 1733
      : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1290 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1211
OY 1734 TCTGAGCAATTAAGTAAAGACACAGTGGATTTATCTAGAGTCAATTAAGTCTGTAG 1793
      : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1230 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1171
OY 1794 TGTGCTTACTCTGAAAAGTGGTGTACATGCTACTGAACCATCAACGAGCATTTAT 1853
      : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1170 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1111
OY 1854 TAAAGACTCTTAAGCAAGACAGATGATGTGAGATTCAGATTTCTTATGC 1906
      : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1110 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1038
```

RESULT 2

US-09-784-316-3
; Sequence 3, Application US/09784316
; Patent No. 6461843

```
;; GENERAL INFORMATION:
;; APPLICANT: WEI, Ming-Hui et al.
;; TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC
;; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES
;; FILE REFERENCE: C1001139
;; CURRENT APPLICATION NUMBER: US/09/784,316
;; CURRENT FILING DATE: 2001-02-16
;; NUMBER OF SEQ ID NOS: 5
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 3
;; LENGTH: 65042
;; TYPE: DNA
;; ORGANISM: Human
;; FEATURE:
;; NAME/KEY: misc_feature
;; LOCATION: (1)..(65042)
;; OTHER INFORMATION: n = A,T,C or G
US-09-784-316-3
```

Query Match 1.5%; Score 37; DB 4; Length 65042;
Best Local Similarity 53.0%; Pred. No. 3.3;

Matches 79; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

```
OY 2164 TGTATTACTCAGCGTTTCATCAATTCACCTTGAGTAACAACATGTTTCACTGCA 2223
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 44535 TTTTTCATCTTATTTCTGTTATTCAGATTG3GTAATTTAATGTTCTATTTTCCA 44594
OY 2224 ATTAATAGTCGCGACCAACCTGGTGTAGTATAGTGTGATTAAGACTGGGTCG 2283
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 44595 GTTATGATGATTTCTCTGCTGCTGCCGTCATCTGTTGTGAGGCTATCACTAGCTTT 44654
OY 2284 TAAATTTCTGCTGAAGCATTTTATATAC 2312
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 44655 TTATTTTGTATGTAATTTTAAATTC 44683
```

RESULT 3

US-08-961-527-113
; Sequence 113, Application US/08961527
; Patent No. 6420135

```
;; GENERAL INFORMATION:
;; APPLICANT: Charles Kunsch
;; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
;; NUMBER OF SEQUENCES: 391
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Human Genome Sciences, Inc.
;; STREET: 9410 Key West Avenue
;; CITY: Rockville
;; STATE: Maryland
;; COUNTRY: USA
;; ZIP: 20850
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage
;; COMPUTER: HP Vectra 486/33
;; OPERATING SYSTEM: MSDOS version 6.2
;; SOFTWARE: ASCII Text
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/961,527
;; FILING DATE:
;; CLASSIFICATION: 424
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER:
;; FILING DATE:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Brookes, A. Anders
;; REGISTRATION NUMBER: 36,373
;; REFERENCE/DOCKET NUMBER: PB340P1
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (301) 309-8504
;; TELEFAX: (301) 309-8512
;; INFORMATION FOR SEQ ID NO: 113:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 18627 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: double
;; TOPOLOGY: linear
US-08-961-527-113
```

Query Match 1.5%; Score 35.2; DB 4; Length 18627;
Best Local Similarity 58.7%; Pred. No. 5.6;
Matches 61; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

```
OY 1213 GTACCAAAACTGAGCTTGTCTTTCAGACATGACAAACAGCTGGTTCAGCTCTG 1272
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 15487 GCAGCAACTGTTTCTTACGTGGTCTGTAGCCATGCGACAGCATGCTGAGGAATG 15546
OY 1273 TTACATTATCTTGAGTGGTGTGACCTATCTGATCCATGTAAGA 1316
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 15547 GTGCCACCACTTGCAATCTTGTGCAACATCTTCTTTTCAAGA 15590
```

RESULT 4

US-09-426-290-1/c
; Sequence 1, Application US/09426290
; Patent No. 6410712

```
;; GENERAL INFORMATION:
;; APPLICANT: Berglund Ran Olafsdottir
;; APPLICANT: Jeffrey Gulcher
;; TITLE OF INVENTION: HUMAN NARCOLEPSY GENE
;; FILE REFERENCE: 2345,2001-000
;; CURRENT APPLICATION NUMBER: US/09/426,290
;; CURRENT FILING DATE: 1999-10-25
;; NUMBER OF SEQ ID NOS: 24
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 1
;; LENGTH: 168575
;; TYPE: DNA
;; ORGANISM: Homo Sapiens
;; FEATURE:
```


NAME/KEY: CDS
LOCATION: (2181)...(21403)
NAME/KEY: CDS
LOCATION: (95252)...(95430)
NAME/KEY: CDS
LOCATION: (101753)...(101996)
NAME/KEY: CDS
LOCATION: (110324)...(110439)
NAME/KEY: CDS
LOCATION: (124058)...(124278)
NAME/KEY: CDS
LOCATION: (127009)...(127130)
NAME/KEY: CDS
LOCATION: (128910)...(129139)
US-09-426-290-1

Query Match 1.5%: Score 35.2; DB 4; Length 168575;
Best Local Similarity 52.0%; Pred. No. 21;
Matches 79; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

QY 2245 TTGGTGTAGTATCTGTTTGCATATAGACTGGGTCTAAATTTCTGGTGAAGCAT 2304
DB 79162 TTGATATATTAACCTGATCTGCATCAGACATTTGATTAACATTTCTAGTTTGGGCT 79103
QY 2305 TTTATACCATTTGTAAGCTTTTACTCTTGAAACTTGGGGAATAAATTAAGT 2364
DB 79102 CACACTGAGTGAAGTGAAGATGATATCTTTAGACACATGTCGGAAGTATGATCA 79043
QY 2365 TGATTTCAATCTCTCAAAAAA 2396
DB 79042 TAATTTGACCTCTATCAAGAACACACA 79011

RESULT 5
US-09-676-610B-24

Sequence 24, Application US/09676610B
Patent No. 6444465
GENERAL INFORMATION:
APPLICANT: C. Frank Bennett
APPLICANT: Jacqueline Wyatt
TITLE OF INVENTION: OLIGONUCLEOTIDE INHIBITION OF HER-1 EXPRESSION
FILE REFERENCE: RTS-0138
CURRENT APPLICATION NUMBER: US/09/676,610B
CURRENT FILING DATE: 2000-09-29
NUMBER OF SEQ ID NOS: 182
SEQ ID NO 24

LENGTH: 169998
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: exon
LOCATION: (1208)...(1472)
NAME/KEY: intron
LOCATION: (1473)...(124390)
NAME/KEY: exon
LOCATION: (124391)...(124544)
NAME/KEY: intron
LOCATION: (124545)...(125409)
NAME/KEY: exon
LOCATION: (125410)...(125595)
NAME/KEY: intron
LOCATION: (125596)...(128711)
NAME/KEY: exon
LOCATION: (128712)...(128848)
NAME/KEY: intron
LOCATION: (128849)...(133400)
NAME/KEY: exon
LOCATION: (133401)...(133469)
NAME/KEY: intron
LOCATION: (133470)...(134652)
NAME/KEY: exon
LOCATION: (134653)...(134773)

NAME/KEY: intron
LOCATION: (134774)...(136116)
NAME/KEY: exon
LOCATION: (136117)...(136261)
NAME/KEY: intron
LOCATION: (136262)...(137936)
NAME/KEY: exon
LOCATION: (137937)...(138053)
NAME/KEY: intron
LOCATION: (138054)...(138637)
NAME/KEY: exon
LOCATION: (138638)...(138766)
NAME/KEY: intron
LOCATION: (138767)...(138864)
NAME/KEY: exon
LOCATION: (138865)...(138940)
NAME/KEY: intron
LOCATION: (138941)...(139765)
NAME/KEY: exon
LOCATION: (139766)...(139860)
NAME/KEY: intron
LOCATION: (139861)...(142245)
NAME/KEY: exon
LOCATION: (142246)...(142445)
NAME/KEY: intron
LOCATION: (142446)...(143605)
NAME/KEY: exon
LOCATION: (143606)...(143738)
NAME/KEY: intron
LOCATION: (143739)...(145838)
NAME/KEY: exon
LOCATION: (145839)...(145931)
NAME/KEY: intron
LOCATION: (145932)...(147385)
NAME/KEY: exon
LOCATION: (147386)...(147544)
NAME/KEY: intron
LOCATION: (147545)...(153274)
NAME/KEY: exon
LOCATION: (153275)...(153321)
NAME/KEY: intron
LOCATION: (153322)...(155088)
NAME/KEY: exon
LOCATION: (155089)...(155231)
NAME/KEY: intron
LOCATION: (155232)...(156025)
NAME/KEY: exon
LOCATION: (156026)...(156151)
NAME/KEY: intron
LOCATION: (156152)...(156826)
NAME/KEY: exon
LOCATION: (156827)...(156928)
NAME/KEY: intron
LOCATION: (156929)...(163399)
NAME/KEY: exon
LOCATION: (163400)...(163586)
US-09-676-610B-24

Query Match 1.4%: Score 34.6; DB 4; Length 169998;
Best Local Similarity 50.3%; Pred. No. 33;
Matches 85; Conservative 0; Mismatches 84; Indels 0; Gaps 0;

QY 843 GAAATGACAGATGCAATTAATTAACCCACAGCTCGACAGAAAGCATGTAACT 902
DB 51241 GAAATTCACATAATGCAATTAACCTAGCAAAATGCTATGTTATTTTCTTAAT 51300
QY 903 GTCACTGATCATTTGACTCTTCAGTCTGTCTGCAAGTGAATTCAGAAACCCCTTGT 962
DB 51301 TCAGATTAATTAAGACATTCATCTCGAAATGCTGTGTCATGTAAGAAAGAAATTTAT 51360
QY 963 AGAGCATGCAATACGATATGCTATAGCTGCTCTATTTGGACACCGAG 1011
DB 51361 CGAGGTGGCCCTTGAGTGCAACAGCCTGTCTCAGCTGCAAAATGAG 51409

RESULT 6
US-09-552-322-5
; Sequence 5, Application US/09552322
; Patent No. 6436642
; GENERAL INFORMATION:
; APPLICANT: Gould-Rothberg
; APPLICANT: Rastelli
; TITLE OF INVENTION: METHOD OF CLASSIFYING A THYROID CARCINOMA USING
; TITLE OF INVENTION: DIFFERENTIAL GENE EXPRESSION
; FILE REFERENCE: 15966-548
; CURRENT APPLICATION NUMBER: US/09/552,322
; CURRENT FILING DATE: 2000-04-19
; PRIOR APPLICATION NUMBER: 60/130,123
; PRIOR FILING DATE: 1999-04-20
; PRIOR APPLICATION NUMBER: 60/193,203
; PRIOR FILING DATE: 2000-03-30
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 5
; LENGTH: 1896
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Assembled
; OTHER INFORMATION: using sequences from AC024267
; NAME/KEY: misc.feature
; LOCATION: (25)..(1291)
; OTHER INFORMATION: wherein any n is a, g, c, or t.
US-09-552-322-5

Query Match 1.4%; Score 34.4; DB 4; Length 1896;
Best Local Similarity 59.0%; Pred. No. 2.5;
Matches 59; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

Qy 2297 GTAGCATTTTATACCATGTGAGTCTTACTCTTGGAACATTCGGGAAATAA 2356
Db 1796 GTACCAATGACATGATGATGATGATGATGATGATGATGATGATGATG 1855
Qy 2357 AATAAGTGTATTCATATCTCTCAAAAAAAAAAAAAAAAAA 2396
Db 1856 TATATATAAACAATAATCTCTCAAAAAAAAAAAAAAAAAA 1895

RESULT 7
US-07-991-867B-25/C
; Sequence 25, Application US/07991867B
; Patent No. 5476781
; GENERAL INFORMATION:
; APPLICANT: Moyer, Richard W.
; APPLICANT: Hall, Richard L.
; APPLICANT: Grudl, Michael E.
; TITLE OF INVENTION: No. 5476781e1 Entomopoxvirus Expression System
; NUMBER OF SEQUENCES: 66
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David R. Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: USA
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/991,867B
; FILING DATE: 12-DEC-1992
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO 92/14818

FILING DATE: 12-FEB-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/827,685
; FILING DATE: 30-JAN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/657,584
; FILING DATE: 19-FEB-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: UF114.C3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 904-375-8100
; TELEFAX: 904-372-5800
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1395 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
US-07-991-867B-25

Query Match 1.4%; Score 34.2; DB 1; Length 1395;
Best Local Similarity 56.8%; Pred. No. 2.3;
Matches 63; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

Qy 1936 ATAGACCTGATCTCAACATCACTTTTGATCTCAATGAGTCTGAGGATTTCTA 1995
Db 291 ATAAATCTAATTTGATTTAAATAATTAATTTATCTGATGATTTATTTGGGATTTA 232
Qy 1996 CTGCTGCTGCTCTCGAGCTGTGATTTGATGATGATTTGATTTGATTTGATTT 2046
Db 231 TTGCTAGTAGTAGAGGATGATTTATTTATTTCTGATATGATATTAAT 181

RESULT 8
US-08-107-755A-25/C
; Sequence 25, Application US/08107755A
; Patent No. 5721352
; GENERAL INFORMATION:
; APPLICANT: Moyer, Richard W.
; APPLICANT: Hall, Richard L.
; APPLICANT: Grudl, Michael E.
; TITLE OF INVENTION: No. 5721352e1 Entomopoxvirus Expression System
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David R. Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: Florida
; COUNTRY: U.S.A.
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/107,755A
; FILING DATE: 19-AUG-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/827,658
; FILING DATE: 30-JAN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/657,584
; FILING DATE: 19-FEB-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: UF114.C2
; TELECOMMUNICATION INFORMATION:

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

TITLE OF INVENTION: NO 5721352e] Entomovirus Expression System

1 NUMBER OF SEQUENCES: 40
 2 CORRESPONDENCE ADDRESS:
 3 ADDRESSEE: David R. Saliwanchik
 4 STREET: 2421 N.W. 41st Street, Suite A-1
 5 CITY: Gainesville
 6 STATE: Florida
 7 COUNTRY: U.S.A.
 8 ZIP: 32606
 9
 10 COMPUTER READABLE FORM:
 11 MEDIUM TYPE: floppy disk
 12 COMPUTER: IBM PC compatible
 13 OPERATING SYSTEM: PC-DOS/MS-DOS
 14 SOFTWARE: PatentIn Release #1.0, Version #1.25
 15
 16 CURRENT APPLICATION DATA:
 17 APPLICATION NUMBER: US/08/107,755A
 18 FILING DATE: 19-AUG-1993
 19 CLASSIFICATION: 435
 20
 21 PRIOR APPLICATION DATA:
 22 APPLICATION NUMBER: US 07/827,658
 23 FILING DATE: 30-JAN-1992
 24
 25 PRIOR APPLICATION DATA:
 26 APPLICATION NUMBER: US 07/657,584
 27 FILING DATE: 19-FEB-1991
 28
 29 ATTORNEY/AGENT INFORMATION:
 30 NAME: Saliwanchik, David R.
 31 REGISTRATION NUMBER: 31,794
 32
 33 REFERENCE/DOCKET NUMBER: 0F114.C2
 34
 35 TELECOMMUNICATION INFORMATION:
 36 TELEPHONE: (904) 375-8100
 37 TELEFAX: (904) 372-5800
 38
 39 INFORMATION FOR SEQ ID NO: 1:
 40
 41 SEQUENCE CHARACTERISTICS:
 42 LENGTH: 6768 base pairs
 43 TYPE: nucleic acid
 44 STRANDEDNESS: double
 45
 46 TOPOLOGY: unknown
 47
 48 MOLECULE TYPE: DNA (genomic)
 49 ORIGINAL SOURCE:
 50 ORGANISM: Amsacta moorei entomopoxvirus
 51
 52 FEATURE:
 53 NAME/KEY: CDS
 54 LOCATION: complement (65..1459)
 55
 56 FEATURE:
 57 NAME/KEY: CDS
 58 LOCATION: 1474..2151
 59
 60 FEATURE:
 61 NAME/KEY: CDS
 62 LOCATION: complement (2239..2475)
 63
 64 FEATURE:
 65 NAME/KEY: CDS
 66 LOCATION: 2502..2987
 67
 68 FEATURE:
 69 NAME/KEY: CDS
 70 LOCATION: 3080..6091
 71
 72 FEATURE:
 73 NAME/KEY: CDS
 74 LOCATION: complement (6277..6768)
 75
 76 US-08-107-755A-1

[illegible]

US-07-991-867B-1/C
Sequence 1, Application US/07991867B
Patent No. 5476781
GENERAL INFORMATION:
APPLICANT: Moyer, Richard W.
APPLICANT: Hall, Richard L.
APPLICANT: Gruidl, Michael E.
TITLE OF INVENTION: No. 5476781el Entomopoxvirus Expression System
NUMBER OF SEQUENCES: 66
CORRESPONDENCE ADDRESS:
ADDRESSEE: David R. Saltwanchik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/991,867B
FILING DATE: 12-DEC-1992
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO 92/14818
FILING DATE: 12-FEB-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/827,685
FILING DATE: 30-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/657,584
FILING DATE: 19-FEB-1991
ATTORNEY/AGENT INFORMATION:
NAME: Saltwanchik, David R.
REGISTRATION NUMBER: 31,794
REFERENCE/DOCKET NUMBER: UF114.C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 904-375-8100
TELEFAX: 904-372-5600
INFORMATION FOR SEQ. ID NO. 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 8457 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: Amsacta moorei entomopoxvirus
FEATURE:
NAME/KEY: CDS
LOCATION: complement (65..1459)
FEATURE:
NAME/KEY: CDS
LOCATION: 1474..2151
FEATURE:
NAME/KEY: CDS
LOCATION: complement (2239..2475)
FEATURE:
NAME/KEY: CDS
LOCATION: 2502..2987
FEATURE:
NAME/KEY: CDS
LOCATION: 3080..6091
FEATURE:
NAME/KEY: CDS
LOCATION: complement (6277..6768)
US-07-991-867B-1

Query Match	1.48;	Score 34.2;	DB 1;	Length 8457;
Best Local Similarity	56.88;	Pred. No. 7;		
Matches 63;	Conservative 0;	Mismatches 48;	Indels 0;	Gaps 0;

```

Oy      1936 ATGAGCGTCGAATCTCAACTCATTGGACATTTGCACATGAGCGTGAAGAAGACTGTTA    1995
        ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db       355 AATAAAATCTAAAATTTTGAAATTTAAAAAATTATATATTCTCATSAGTTTTATTTGGCATTTA    296

Oy      1996 CTTCGCTGCTCCCTCGAGCGTGTGGATTTGATGCCAFATTTGCACATCACT    2046
        + | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db       295 TTGCGTAGTAGTTATAGAGGTGCATTAATTATTCTTAGATGATGATTAAT    245


RESULT 13
US-08-544 -332-1/c
; Sequence 1, Application US/08544332
Patent No. 5935777

GENERAL INFORMATION:
APPLICANT: Moyer, Richard W.
APPLICANT: Hall, Richard L.
APPLICANT: Gruidl, Michael E.
TITLE OF INVENTION: No. 5935777e1 Entomopoxvirus Expression System
NUMBER OF SEQUENCES: 77
CORRESPONDENCE ADDRESS:
ADDRESSEE: Gerard H. Bencen
STREET: 2421 N.W. 41st Street, Suite A-1
City: Gainesville ,
STATE: FL
COUNTRY: USA
ZIP: 32606

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/544,332
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/991,867
FILING DATE: 07-DEC-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/107,755
FILING DATE: 19-AUG-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO 92/14818
FILING DATE: 12-FEB-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/827,685
FILING DATE: 30-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/657,584
FILING DATE: 19-FEB-1991
ATTORNEY/AGENT INFORMATION: .
NAME: Benven, Gerard H.
REGISTRATION NUMBER: 35,746
REFERENCE/DOCKET NUMBER: UF114.C4
TELECOMMUNICATION INFORMATION:
TELEPHONE: 904-375-5800
TELEFAX: 904-372-5800
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 8457 base pairs
TYPE: nucleic acid
STRANDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: Amsacta moorei entomopoxvirus
FEATURE:
NAME/KEY : CDS
LOCATION: complement (65..1459)
FEATURE:
NAME/KEY : CDS
LOCATION: 1474..2151
```

```

FEATURE: NAME/KEY: CDS LOCATION: complement (2239..2475)
FEATURE: NAME/KEY: CDS LOCATION: 2502..2987
FEATURE: NAME/KEY: CDS LOCATION: 3080..6091
FEATURE: NAME/KEY: CDS LOCATION: complement (6277..6768)
US-08-544-332-1

Query Match Best Local Similarity 1.4%; Score 34.2; DB 4; Length 8457;
Matches 63; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

QY 1936 ATAGAGCCGTGATCTCAACCTCATCTTTGGATTCTACAATAGTGCAGAGACTTCTA 1995
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 355 ATAAATCTAAATTGAATTTAAAAAATTAATATTCATAGAGCTTTATTTGGTGATTTA 296

QY 1996 CTGCGTGTTGCCGAGCTGGTGATTTGATGCAATATTTGCATCACT 2046
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 295 TTGCCTAGTAGTTATAGAGGTGCATTAATTTCTAGATATGATAATTAAT 245

RESULT 14
US-09-370-861A-1/c
Sequence 1, Application US/09370861A
Patent No. 6410221
GENERAL INFORMATION:
APPLICANT: Moyer, Richard W.
APPLICANT: Hall, Richard L.
APPLICANT: Gruidl, Michael E.
TITLE OF INVENTION: No. 6410221el Entomopoxvirus Expression System
FILE REFERENCE: UPI14.C4.D1
CURRENT APPLICATION NUMBER: US/09/370.861A
CURRENT FILING DATE: 1999-08-09
PRIOR APPLICATION NUMBER: US 07/991,867
PRIOR FILING DATE: 1992-12-07
PRIOR APPLICATION NUMBER: US 08/107,755
PRIOR FILING DATE: 1993-08-19
PRIOR APPLICATION NUMBER: WO 92/14818
PRIOR FILING DATE: 1992-02-12
PRIOR APPLICATION NUMBER: US 07/827,685
PRIOR FILING DATE: 1992-01-30
PRIOR APPLICATION NUMBER: US 07/657,584
PRIOR FILING DATE: 1991-02-19
NUMBER OF SEQ ID NOS: 78
SOFTWARE: PatentIn version 3.1
SEQ ID NO 1
LENGTH: 8457
TYPE: DNA
ORGANISM: Amsacta moorei entomopoxvirus
US-09-370-861A-1

Query Match Best Local Similarity 1.4%; Score 34.2; DB 4; Length 8457;
Matches 63; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

QY 1936 ATAGAGCCGTGATCTCAACCTCATCTTTGGATTCTACAATAGTGCAGAGACTTCTA 1995
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 355 ATAAATCTAAATTGAATTTAAAAAATTAATATTCATAGAGCTTTATTTGGTGATTTA 296

QY 1996 CTGCGTGTTGCCGAGCTGGTGATTTGATGCAATATTTGCATCACT 2046
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 295 TTGCCTAGTAGTTATAGAGGTGCATTAATTTCTAGATATGATAATTAAT 245

RESULT 15
US-09-253-316-3/c
Sequence 3, Application US/09253316

```

```

Patent NO. 6395890
GENERAL INFORMATION:
APPLICANT: Sheppard, Paul O.
APPLICANT: Jaspers, Stephen R.
TITLE OF INVENTION: CONNECTIVE TISSUE GROWTH FACTOR HOMOLOGS
FILE REFERENCE: 97-75
CURRENT APPLICATION NUMBER: US/09/253,316
CURRENT FILING DATE: 1999-02-19
EARLIER APPLICATION NUMBER: US 60/075,300
EARLIER FILING DATE: 1998-02-20
NUMBER OF SEQ ID NOS: 34
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 3
LENGTH: 1062
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Degenerate sequence
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(1062)
OTHER INFORMATION: n = A,T,C or G
US-09-253-316-3

```

GenCore version 5.1.4.p5.4578.
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 26, 2003, 12:53:20 ; Search time 88 Seconds
(without alignments)
1182.429 Million cell updates/sec

Title: US-09-988-863A-2
Perfect score: 2581
Sequence: 1 MAVVASAPKVKLTGCVLYL.....ESGDPRTCTGCVSIHLE 505

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: SPREMBL_21:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_protist:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rviro:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	ID	Description
1	2581	100.0	505 10 Q9C6T1	Q9C6T1 arabidopsis
2	1930.5	74.8	503 10 Q944G1	Q944G1 hevea bras
3	493	19.1	426 3 Q9UT88	Q9UT88 schizosacch
4	235	9.1	314 17 Q973B5	Q973B5 sulfolobus
5	230	8.9	323 17 Q97UL6	Q97UL6 sulfolobus
6	178.5	6.9	368 2 Q9FDE7	Q9FDE7 enterococcu
7	158.5	6.1	294 2 Q9K3K1	Q9K3K1 lactobacilli
8	155	6.0	374 2 Q9KWC3	Q9KWC3 streptomyce
9	151.5	5.9	360 16 Q92F01	Q92F01 listeria in
10	149.5	5.8	359 16 Q8YAV1	Q8YAV1 listeria mo
11	147.5	5.7	361 2 Q9FDE2	Q9FDE2 enterococcu
12	142.5	5.5	358 2 Q9FDE3	Q9FDE3 enterococcu
13	136.5	5.3	360 2 Q9KWF7	Q9KWF7 staphylococ
14	136.5	5.3	360 2 Q9KWF7	Q9KWF7 kitasatospo
15	134.5	5.2	358 16 Q99W20	Q99W20 staphylococ
16	130.5	5.1	358 2 Q9FDT7	Q9FDT7 staphylococ

17	130.5	5.1	434 11 Q9D307	Q9D307 mus musculu
18	130	5.0	816 5 Q62219	Q62219 caenorhabdi
19	126.5	4.9	241 11 Q9D0B4	Q9D0B4 mus musculu
20	125	4.8	376 10 Q9STB1	Q9STB1 arabidopsis
21	125	4.8	816 2 Q9Z1L9	Q9Z1L9 lactococcus
22	124.5	4.8	317 16 Q51630	Q51630 borrellia bu
23	123.5	4.6	352 5 Q8SUB3	Q8SUB3 encyphalito
24	118.5	4.6	334 17 Q8U0F3	Q8U0F3 pyrococcus
25	117.5	4.6	721 16 Q8VM05	Q8VM05 anabaena sp
26	117	4.5	386 10 Q94AG2	Q94AG2 hevea bras
27	115.5	4.5	4625 10 Q9SMH3	Q9SMH3 chlamydomon
28	115	4.5	816 16 Q9CHS9	Q9CHS9 lactococcus
29	114	4.4	823 5 Q9VMB0	Q9VMB0 diosiphilla
30	114	4.4	3919 16 Q9CPH9	Q9CPH9 pasteurella
31	112.5	4.4	478 16 Q8UHP9	Q8UHP9 agrobacteri
32	112.5	4.4	902 16 Q91742	Q91742 pseudomonas
33	112	4.3	797 17 Q8TPP4	Q8TPP4 methanosarc
34	110.5	4.3	1417 16 Q9HWR8	Q9HWR8 pseudomonas
35	109.5	4.2	1101 4 Q9NXY9	Q9NXY9 homo sapien
36	109.5	4.2	1348 11 Q08828	Q08828 mus musculu
37	109	4.2	3475 5 Q9U175	Q9U175 leishmania
38	108.5	4.2	615 16 Q8UJA0	Q8UJA0 agrobacteri
39	108.5	4.2	659 16 Q8ZHC8	Q8ZHC8 yersinia pe
40	108.5	4.2	1477 2 Q48028	Q48028 haemophilus
41	107	4.1	463 16 Q9S243	Q9S243 streptomyce
42	106.5	4.1	715 2 Q30965	Q30965 rhizobium 1
43	106	4.1	382 10 Q94B09	Q94B09 arabidopsis
44	106	4.1	667 12 Q67482	Q67482 japanese en
45	106	4.1	1087 10 Q82303	Q82303 arabidopsis

ALIGNMENTS

RESULT 1
ID Q9C6T1 PRELIMINARY; PRT; 505 AA.
AC Q9C6T1;
DT 01-JUN-2001 (TREMBLrel. 17, last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, last annotation update)
DE Hypothetical 54.4 kDa protein.
OS F5M6.9.
GN Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC Eucosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV COLUMBIA;
RX MEDLINE=21016719; PubMed=11130712;
RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
RA White O., Alonso J., Altati H., Araujo R., Bowman C.L., Brooks S.Y.,
RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
RA Dunn P., Etgu P., Feldlyum T.V., Feng J.-D., Fong B., Fujii C.Y.,
RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,
RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
RA Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,
RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
RA Lin X., Liu S.X., Liu Z.A., Lueros J.S., Maiti R., Marziani A.,
RA Maltischer J., Miranda M., Nguyen M., Nielsen M.C., Osborne B.I.,
RA Pal G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
RA Utecherack T., Van Aken S., Vaysberg M., Vysotskaya V.S., Walker M.,
RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.,
RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
thaliana.";
RL Nature 408:816-820(2000).
DR EMBL; AC079041; AAG50716.1;
TIGRFAMS; TIGR01219; Pmev_kin_ERG8; 1.

KW Hypothetical protein.
 SQ SEQUENCE 505 AA; 54409 MW; BICBA6CA338B3D63 CRC64;
 Query Match 100.0%; Score 2581; DB 10; Length 505;
 Best Local Similarity 100.0%; Pred. No. 1.9e-190;
 Matches 505; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAVVASAPGKVLMTGGLYLEKPNAGLVLTSTNARFAIVKPIINEVKEPESAMKMTDVKL 60
 DB 1 MAVVASAPGKVLMTGGLYLEKPNAGLVLTSTNARFAIVKPIINEVKEPESAMKMTDVKL 60
 QY 61 TSPQLRESMYKLSLNHLTQSVASASDRNPFEVHAIOYIAAAHLATEKDKESLHKL 120
 DB 61 TSPQLRESMYKLSLNHLTQSVASASDRNPFEVHAIOYIAAAHLATEKDKESLHKL 120
 QY 121 QGDDITILGNDPFSYRNQIESAGLPLTPESLGLTAPFASITFNAESNGANSKPEVAKT 180
 DB 121 QGDDITILGNDPFSYRNQIESAGLPLTPESLGLTAPFASITFNAESNGANSKPEVAKT 180
 QY 181 GLGSSAAMTAVVAALLHLYGVVDLSDPCKEKGKFGCSLDVYIHMIAQTSCLAQKVGSG 240
 DB 181 GLGSSAAMTAVVAALLHLYGVVDLSDPCKEKGKFGCSLDVYIHMIAQTSCLAQKVGSG 240
 QY 241 FDVSCAVYGSORVYRSPVLSFAQVAVGGLPNEVYIGTILGKMDKKRTEFSLPLMLN 300
 DB 241 FDVSCAVYGSORVYRSPVLSFAQVAVGGLPNEVYIGTILGKMDKKRTEFSLPLMLN 300
 QY 301 FLGEPGSGGSTPSMVGAVKKMQMDPEKARENQNLSDANLELETKLNDLSKLAKDHW 360
 DB 301 FLGEPGSGGSTPSMVGAVKKMQMDPEKARENQNLSDANLELETKLNDLSKLAKDHW 360
 QY 361 VYLKRVKSCSVLTSEKWLHATEPINEAIIKELLEAREAMLRIIRILROMGEAASVPIE 420
 DB 361 VYLKRVKSCSVLTSEKWLHATEPINEAIIKELLEAREAMLRIIRILROMGEAASVPIE 420
 QY 421 ESOTOLDSTMSAEGVLLAGVPGAGFDPAIFATLIGDSGKTLQAWSSHNVALLVREDP 480
 DB 421 ESOTOLDSTMSAEGVLLAGVPGAGFDPAIFATLIGDSGKTLQAWSSHNVALLVREDP 480
 QY 481 HGVCLESQDPRITCITSGVSSIHLE 505
 DB 481 HGVCLESQDPRITCITSGVSSIHLE 505

RESULT 2
 Q944G1 PRELIMINARY; PRT; 503 AA.
 AC 0944G1:
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Phosphomevalonate kinase.
 OS Hevea brasiliensis (Para rubber tree).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids I; Malpighiales; Euphorbiaceae; Hevea.
 NCBI_TaxID=3981;
 RN
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. RIM600;
 RA Hallahan D.L., Kelper-Hyenko N.M.;
 RT "Genes involved in the biosynthesis of isopentenyl diphosphate in the
 rubber tree *Hevea brasiliensis*."
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF429385; AAL18926.1;
 DR TIGRfams; TIGR01219; Pmev_kin_ERG8; 1.
 KW kinase.
 SQ SEQUENCE 503 AA; 54171 MW; BE6F91B80B45FF94 CRC64;

Query Match 74.8%; Score 1930.5; DB 10; Length 503;
 Best Local Similarity 74.1%; Pred. No. 2.7e-140;
 Matches 374; Conservative 61; Mismatches 67; Indels 3; Gaps 2;

QY 1 MAVVASAPGKVLMTGGLYLEKPNAGLVLTSTNARFAIVKPIINEVKEPESAMKMTDVKL 60
 DB 1 MAVVASAPGKVLMTGGLYLEKPNAGLVLTSTNARFAIVKPIINEVKEPESAMKMTDVKL 60
 QY 61 TSPQLRESMYKLSLNHLTQSVASASDRNPFEVHAIOYIAAAHLATEKDKESLHKL 119
 DB 61 TSPQLRESMYKLSLNHLTQSVASASDRNPFEVHAIOYIAAAHLATEKDKESLHKL 120
 QY 120 LOGDDITILGNDPFSYRNQIESAGLPLTPESLGLTAPFASITFNAESNGANSKPEVAK 179
 DB 120 LOGDDITILGNDPFSYRNQIESAGLPLTPESLGLTAPFASITFNAESNGANSKPEVAK 180
 QY 180 TGLSSAAMTAVVAALLHLYGVVDLSDPCKEKGKFGCSLDVYIHMIAQTSCLAQKVGSG 239
 DB 180 TGLSSAAMTAVVAALLHLYGVVDLSDPCKEKGKFGCSLDVYIHMIAQTSCLAQKVGSG 238
 QY 240 GFDVSCAVYGSORVYRSPVLSFAQVAVGGLPNEVYIGTILGKMDKKRTEFSLPLMLN 299
 DB 239 GFDVSCAVYGSORVYRSPVLSFAQVAVGGLPNEVYIGTILGKMDKKRTEFSLPLMLN 298
 QY 300 FLGEPGSGGSTPSMVGAVKKMQMDPEKARENQNLSDANLELETKLNDLSKLAKDHW 359
 DB 299 FLGEPGSGGSTPSMVGAVKKMQMDPEKARENQNLSDANLELETKLNDLSKLAKDHW 358
 QY 360 VYLKRVKSCSVLTSEKWLHATEPINEAIIKELLEAREAMLRIIRILROMGEAASVPIE 419
 DB 359 VYLKRVKSCSVLTSEKWLHATEPINEAIIKELLEAREAMLRIIRILROMGEAASVPIE 418
 QY 420 PESOTOLDSTMSAEGVLLAGVPGAGFDPAIFATLIGDSGKTLQAWSSHNVALLVRED 479
 DB 419 PESOTOLDSTMSAEGVLLAGVPGAGFDPAIFATLIGDSGKTLQAWSSHNVALLVRED 478
 QY 480 PHGVCLESQDPRITCITSGVSSIHLE 504
 DB 479 PHGVCLESQDPRITCITSGVSSIHLE 503

RESULT 3
 Q9UT88 PRELIMINARY; PRT; 426 AA.
 AC 09UT88:
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Putative phosphomevalonate kinase.
 GN SPAC343.01C.
 OS Schizosaccharomyces pombe (Fission Yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomyces.
 NCBI_TaxID=4896;
 RN
 RP SEQUENCE FROM N.A.
 RC STRAIN=972H-;
 RA Murphy L., Harris D., Wood V., Rajandream M.A., Barrell B.G.;
 RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL109739; CAB52264.1;
 DR TIGRfams; TIGR01219; Pmev_kin_ERG8; 1.
 KW kinase.
 SQ SEQUENCE 426 AA; 47322 MW; 25635B99E7FAAD64 CRC64;

Query Match 19.1%; Score 493; DB 3; Length 426;
 Best Local Similarity 30.2%; Pred. No. 1.3e-29;
 Matches 153; Conservative 83; Mismatches 170; Indels 100; Gaps 21;

QY 1 MAVVASAPGKVLMTGGLYLEKPNAGLVLTSTNARFAIVKPIINEVKEPESAMKMTDVKL 60
 DB 1 MAVVASAPGKVLMTGGLYLEKPNAGLVLTSTNARFAIVKPIINEVKEPESAMKMTDVKL 60
 QY 61 TSPQLRESMYKLSLNHLTQSVASASDRNPFEVHAIOYIAAAHLATEKDKE 113
 DB 52 KSPQFINAEMLYINIDWVSPIRVHQIYENCLEKNRPVQALFYI-NYFSTGR--- 107

QY 114 SLKLLQGLDITLLGSNDYFYNQIESAGLPLTPESLGLTAD---FASTFENAASNG 170
 Db 108 --OPLCGMODLQVTLQVNNAY--HQPO-----LKPOQ--TSYPKFNFLNCTLG----- 149
 QY 171 ANSRPEAKTGCGSSAMTAFAVAALHLYGLVNDLSDPC--KEKFGCSDLDVYHMTAQT 228
 Db 150 ----GVHKTGGLGSSAMTSLIGSL--FLSLRLTDTGDKSLKIDSTKVYIHNLAQI 202
 QY 229 SHCLAQGVSGFVSCAVGVNQRYRFPSEVLISFAQV---AATGLPLNEVIGTILKKG 284
 Db 203 AHCSAQGVSGFVSCAVGVNQRYRFPSEVLISFAQV---AATGLPLNEVIGTILKKG 262
 QY 285 WDNKRFESLPLMLNLFLEPGSGSSTPSWGAVKRQMSDPEKARENMONLSDANLEL 344
 Db 263 WSDV--VFQJLPATYCLIMGDV--AGGSTPGWVKVQOQKQENPESEK---NCFD----- 311
 QY 345 ETKLNDLSKLAKHMDVLYR--IKSCSVLTSEKVLHATEPINEAIIKELLEAREAMLK 402
 Db 312 -----DLYSRVLSIKNCFL-----SSSLSLSELOSOFRSIRILQOR 347
 QY 403 IRLRMQGEASVPIEPESOTOLDSTMSAEGVLLAGVPGAGGDAIFAITLGDGSG--T 460
 Db 348 ITV-----EAKVDIPLKQTNILNDINEQLPVIGVPGAGGDAFOCLAINHTEIIE 400
 QY 461 KLQAMSSHNVLLALLVREDPHCYCLE 486
 Db 401 NVIKTKMDGVVPMVDVSPAPDGLAVE 426

RESULT 4
 Q973B5 ID 0973B5 PRELIMINARY; PRT; 314 AA.

AC 0973B5;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Hypothetical protein ST0978.
 GN ST0978.
 OS Sulfolobus tokodaii.
 OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
 OC Sulfolobus.
 OX NCBI_TaxID=111955;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=JCM 10545 / 7;
 RX PubMed=11572479;
 RA Kawarabayashi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M.,
 RA Sekine M., Baba S.-T., Ankai A., Kosugi H., Hosoyama A., Fukui S.,
 RA Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Takamiya M., Kato Y.,
 RA Yoshizawa T., Tanaka T., Kudo Y., Yamazaki J., Kushiida N., Oguchi A.,
 RA Aoki K.-I., Masuda S., Yanagii M., Nishimura M., Yamagishi A.,
 RA Oshima T., Kikuchi H.;
 RT "Complete genome sequence of an aerobic thermophilic
 RT Crenarchaeon, Sulfolobus tokodaii strain7.";
 RL DNA Res. 8:123-140(2001).
 DR EMBL: AP000984; BAB65998.1;
 DR InterPro: IPR001745; GHMPkinase_Arp.
 DR Pfam: PF00288; GHMP_kinases; 1.
 DR Hypothetical protein: Complete proteome.
 KW SEQUENCE 314 AA; 35307 MW; 89740213AB16335F CRC64;

Query Match 9.1%; Score 235; DB 17; Length 314;
 Best Local Similarity 23.2%; Pred. No. 6.3e-10;
 Matches 113; Conservative 64; Mismatches 129; Indels 182; Gaps 19;

QY 6 SAPKVLMTGTYLLEKPNAGLVSTNARPAI--VKPINEVYKPESSAMKWTQVLTSP 63
 Db 3 SAPKILMTGYSYV-----FGISHVIAINRKR-----CDIK----- 36
 QY 64 QLSRESMYKLSLNLTLQSVASDSRNPFEVHAIOVAIAAHLATEKDKESLKLQGL 123
 Db 37 -----SSNNF-----IFETTYGTFRDK----- 53

QY 124 DITLLGSNDYFYNQIESAGLPLTPESLGLTADPASITFENAASNGANSPREYAKTGLG 183
 Db 54 -----CNELIESVITVFKKFGSLPPEHYVLFNDKDFQ--IHGR-----KTGLG 95
 QY 184 SSAMTAFAVAALHLYGLVNDLSDPCRKGRKFGCSDLDVYHMTAQTSHLQAGXVSGFDV 243
 Db 96 SSSASTYALTAICYIYYL-----FRNLKDEIYLAQKANIIRQKJSGFDI 142
 QY 244 SCAYVGSORVYRFPSEVLISFAQVAVTGLPLNEVIGTILKGMKDKRFESLPL-----MN 299
 Db 143 ASAYVGSIVYRFP-----YDIEKDVSLPELKLIGYE 174
 QY 300 LFLGEPGSGSSTPSWGAVKRQMSDPEKARENMONLSDANLEETKLANDLSKLAKDW 359
 Db 175 MLGFIQGSFSTVNSVAFIK-----SNNEEFKVKWKYIDEENI--MAIKLILKG----- 224
 QY 360 DYLRVIRKSCSVLTSEKVLHATEPINEAIIKELLEAREAMLRIILRMQGEASVPI 419
 Db 225 -----IEEAI-----EHVRLARFRLGLAKKIVGEIE 252
 QY 420 PESOTOLDSTMSAEGVLLAGVPGAGGDAIFAITLGDGSGTKLQAMSSHNVLLALLVED 479
 Db 253 NEKIRRLIE--MAENDALIALSPGAG--ESVF--VLGKDLKVKKEKENVIVTELKED 307
 QY 480 PHGVLES 487
 Db 308 -EGLRIEA 314

RESULT 5
 Q97UL6 ID 097UL6 PRELIMINARY; PRT; 323 AA.

AC 097UL6;
 DT 01-OCT-2001 (TREMBLrel. 18, Created)
 DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE (phospho) mevalonate kinase, putative.
 GN S502988.
 OS Sulfolobus solfataricus.
 OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
 OC Sulfolobus.
 OX NCBI_TaxID=2287;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 35092 / DSM 1617 / P2;
 RX MEDLINE=21332296; PubMed=11427726;
 RA She Q., Singh R.K., Confalonieri F., Zivanovic Y., Allard G.,
 RA Aweyer M.J., Chan-Weiler C.C.-Y., Clausen I.G., Curtis B.A.,
 RA De Moors A., Eranuso G., Fletcher C., Gordon P.M.K.,
 RA Helkamp-de Jong I., Jeffries A.C., Kozera C.J., Medina N., Peng X.,
 RA Thi-Ngoc H.P., Redder P., Schenk M.E., Thieriault C., Tolstrup N.,
 RA Charlebois R.L., Doellittle W.F., Duguet M., Gaasterland T.,
 RA Garrett R.A., Ragan M.A., Sensen C.W., Van der Oost J.,
 RT "The complete genome of the crenarchaeon Sulfolobus solfataricus P2.";
 RT Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).
 RL EMBL: AE006889; AAK43093.1;
 DR InterPro: IPR001174; Galkkinase.
 DR InterPro: IPR001745; GHMPkinase_Arp.
 DR Pfam: PF00288; GHMP_kinases; 1.
 DR PRINTS: PR00960; LMBPPTOIN.
 KW kinase; Complete proteome.
 KW SEQUENCE 323 AA; 36207 MW; D62E48552F5CBDF8 CRC64;

Query Match 8.9%; Score 230; DB 17; Length 323;
 Best Local Similarity 22.2%; Pred. No. 1.6e-09;
 Matches 110; Conservative 72; Mismatches 129; Indels 184; Gaps 21;

QY 3 VVASAPGKVLMTGTYLLEKPNAGL--VLSTNARPAIYKPKINEVYKPESSAMKWTQVXL 60
 Db 2 IKVASPGKILMTGYSYV-----FGISHVIAVKNRYSLSREIKE----- 42
 QY 61 TSPQLSRESM-YKLSLNLTLQSVASDSRNPFEVHAIOVAIAAHLATEKDKESLKL 119
 Db 119 -----

```

Db 43 -----KSLIFHTSYGHF-----KNSGNEILNSVL-----DTERERLSQ-L 77
OY 120 LOGIDITILGSDNDFYSYRNOIESAGLPLTFESIGTLAPFASITFENNAESNGANSKPREVAK 179
Db 78 PGGEYIDLNDKERLI-----DGGK 97
OY 180 TGLGSSAAMTTAVVAALHLHYGVVLDSDPCKEKGCSLDLVHIMIAQTSCHLAQGVGS 239
Db 98 TGLGSSAAMTVSLTACLYVAI-----HGKTL--DLFEIKRLAQIINRYKQKIGS 144
OY 240 GFDVSCAVYGSORVYRFSPEVLSFAQAVATGLPLNEVIGTILKGMKNKREESLPL- 297
Db 145 GFDIASAVFSGSIYKRT-----DLDRKDFPEKTLNL 176
OY 298 --MNLFLGEPGSGSSPTSMVGAVKKN-QMSDEPKAREMNQULSDANLELETKLNDLSKL 354
Db 177 GNDVLMGFGKSSSET-----VGLVRKFEKSNLDDFKEIMLIDEENV-MAIKLIKIKLNL 231
OY 355 --AKDHVDYLRVYKSCSVLTSEKWLHATEPINEAIIKELLEAREMLRIRILMROMGE 412
Db 232 DEAEVH-----IKGRKYLVNVIARI--VGVKLVSKMEELIKI----- 268
OY 413 AASVPIPESTQTLDDSTMSAEGVLLAGVPGAGFDALFAITLGDSCGTLQAMSSHNVL 472
Db 269 -----AEEEGALVALSPGAGGDSIFA--LGNDLNRVREANSKRGIF 308
OY 473 ALLVREDPHGVCLSS 487
Db 309 IIDVKED-EGRLLES 322

RESULT 6
OY 09FD67 ID 09FD67 PRELIMINARY; PRT; 368 AA.
AC 09FD67;
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Phosphomevalonate kinase.
GN MYAK2.
OS Enterococcus faecalis (Streptococcus faecalis).
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
OC Enterococcaceae; Enterococcus.
OX NCBI_TaxID=1351;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20353468; PubMed=10894743;
RA Wilding E.I., Brown J.R., Bryant A.P., Chalker A.F., Holmes D.J.,
RA Ingraham K.A., Iordanescu S., So C.Y., Rosenberg M., Gwynn M.N.,
RT "Identification, evolution and essentiality of the mevalonate pathway
RT for isopentenyl diphosphate biosynthesis in gram-positive cocci".
RL J. Bacteriol. 182:4319-4327(2000).
DR EMBL; AF290093; AAC02442.1;
DR InterPro; IPR001745; GHMPase_AFP.
DR InterPro; IPR001459; Mev_gal_kin.
DR Pfam; PF00288; GHMP_kinases; 1.
DR PRINTS; PR00959; MEVGALKINASE.
DR TIGRfams; TIGR01220; Pmev_kin_Gr_pos; 1.
KW Kinase.
SQ SEQUENCE 368 AA; 40541 MW; 390993036577146D CRC64;

Query Match 6.9%; Score 178.5; DB 2; Length 368;
Best Local Similarity 22.3%; Pred. No. 1.8e-05;
Matches 106; Conservative 58; Mismatches 179; Indels 133; Gaps 20;
OY 3 VVASAPKVLMTGTYLLEKPNAGLVSTNARFAIVKPIINEEVKPPSMARKTVDKLTLS 62
Db 2 IEVTTPEKFLIAGYAVEEGHPAIIIVADQFVTVTEETDEGSIQSAQSYSLPIKMT- 60
OY 63 POLSRESMYKLSLNLHLQGSASDSNPNFEHAIOYAIAAHLATKDESLHKLILQ 122
Db 61 ----RRNCELVLDIR-----ENPF-----HYVLAIRHL-TEKVAOEONKEL--- 96

```

```

OY 123 LDITILGSDNDFYSYRNOIESAGLPLTFESIGTLAPFASITFENNAESNGANSKPREVATGL 182
Db 97 -----SEYHLK-----VTSSEDSNG-----RRYGL 117
OY 183 GSSAAMTTAVVAAL--LHYGVVLDSDPCKEKGCSLDLVHIMIAQTSCHLAQGVGS 240
Db 118 GSSGAVTVGVKALNIFYDGL-----ENEIRKLSLAHLAVOQN-GSC 161
OY 241 FDVSCAVYGSORVYRFSPEVLSFAQAVATGLPLNEVIGTILKGMKNKRT-EESLPLMN 299
Db 162 GDIAASYCG--WIAFSPDHWNOKVT-----TETLTDLADMPELMIFPLKVPQLR 215
OY 300 LFLGEPGSGSSPTSMVGAVKKNQMSDEPKAREMNQULSDANLELETKLNDLSLADHW 359
Db 216 LLIGWTS-PASTSDLDVRY--HQSKEKQAAVEQFLMSRLCVERMINGFN----- 264
OY 360 DVLVRIKSCSVLTSEKWLHATEPINEAIIKELLEAREMLRIRILMROMGEASVPIE 419
Db 265 -----TGKIS-VIOKQITKRO-----LLAELSLITGVIE 294
OY 420 PESQTQLDDSTMSAEGVLLAGVPGAGFDALFAITLGDSC-TKLQAMSSHNVLAL 474
Db 295 TEALKNLCDLAESYTGA--AKSSGAGGCGGIYIFROKSGILPLMTAMERKGIITPL 348

RESULT 7
OY 093K31 ID 093K31 PRELIMINARY; PRT; 294 AA.
AC 093K31;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Phosphomevalonate kinase.
GN PMK.
OS Lactobacillus helveticus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
OC Lactobacillaceae; Lactobacillus.
OX NCBI_TaxID=1587;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=53/7;
RA Smets A., Purtsi T., Palva A.;
RT "Identification of a gene cluster for the mevalonate pathway in
RT Lactobacillus helveticus".
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ279018; CAC51372.1;
DR InterPro; IPR001745; GHMPase_AFP.
DR Pfam; PF00288; GHMP_kinases; 1.
DR TIGRfams; TIGR01220; Pmev_kin_Gr_pos; 1.
KW Kinase.
SQ SEQUENCE 294 AA; 32537 MW; 0795099A1BC97938 CRC64;

Query Match 6.1%; Score 158.5; DB 2; Length 294;
Best Local Similarity 24.8%; Pred. No. 0.00045;
Matches 75; Conservative 42; Mismatches 117; Indels 69; Gaps 15;
OY 179 KTGIGSSAAMTTAVVAALHLHYGVVLDSDPCKEKGCSLDLVHIMIAQTSCHLAQGVGS 238
Db 49 KYGLGSSAAMTVATVAKILHFGY-VKMSN-----ELVYKLSAISHSYVQGN-G 94
OY 239 GFDVSCAVYGS--ORVYRFSPEVLSFAQAVATGLPLNEVIGTILKGMKNKREESLPL 297
Db 95 SAGDIAASYGVGMVATQTFPKKMLQY-ELA-----NKTLSDVVNEMPGIKELLPLPH 147
OY 298 --MNLFLGEPGSGSSPTSMVGAVKKNQMSDEPKAR-ENMNQULSDANLELETKLNDLSKLA 355
Db 148 DMKLMIG-----WSQKPASTRSLVDETNANAAALNTEYK----- 181
OY 356 KDHDVYLRVYKSCSVLTSEKWLHATEPINEAIIKELLEAREMLRIRILMROMGEAAS 415
Db 182 ----NFLAASSEC-VLK-----MIVGFKAKNIALIKQIIVN-----RHLLQHFARKINQ 225
OY 416 VPIEPESQTQLDDSTMSAEGVLLAGVPGAGFDALFAITLGDSCG-TKLQAMSSHNVLAL 474

```

Db 226 IAIPIPLTQIKIAEDFGA--AKTSGAGGCGIYITDADTDVDALENEMRRNGIPL 283
Qy 475 LVR 477
Db 284 NFR 286

RESULT 8

Q9KMG3 PRELIMINARY; PRT; 374 AA.
AC Q9KMG3;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Phosphomevalonate kinase.
OS Streptomyces sp. (strain CL190).
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=93372;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CL190;
RX MEDLINE=20353446; PubMed=10894721;
RA Takagi M., Kuzuyama T., Takahashi S., Seto H.;
RT "A gene cluster for the mevalonate pathway from Streptomyces sp.
strain CL190."
RL J. Bacteriol. 182:4153-4157(2000).
DR EMBL; AB037666; BAB07792.1;
DR InterPro: IPR001969; Aspartase_site.
DR InterPro: IPR001745; GMPkinase_ATP.
DR Pfam: PF00288; GMP_kinases; 1.
DR PRINTS; PR00959; MEVGAALKINASE.
DR TIGRFAMS; TIGR01220; Pmev_kin_Gr_pos; 1.
DR PROSITE; PS00141; ASP_PROTEASE; UNKNOWN_1.
KW kinase.
SQ SEQUENCE 374 AA; 39281 MW; A2908B33C694E6D CRC64;

Query Match 6.08; Score 155; DB 2; Length 374;

Best Local Similarity 19.78; Pred. No. 0.0012;
Matches 97; Conservative 68; Mismatches 122; Indels 156; Gaps 19;

Qy 3 VVAAPGKVLMTGGLVLEKPNAGLVLTSTNARFYAIK-----INEEYKPSW 51
Db 8 IVRAIPGKLFYAGERYAVDPGNPAIVAVDRIHISVTSADADGADYVSSDGLPGAV 67
Qy 52 AMKWTDVKTSPOLSRSMYKLSLNHLTLQSVASDSRNPVEHAIOYAIANAHLATEKD 111
Db 68 GMRWHDRLV-----VRDP--DDQQAARSALAHVY--- 95
Qy 112 KESLHLKLLQGLDITILGSDNFYSRNOIESAGLPLTPESLGITLAPFASITFNMAESGA 171
Db 96 -----SAIEIVGRILGREGKVPRLTLTVSSRLHEDG- 127
Qy 172 NSKPEVAKTGIGSSAAMTTAVVAALLHYLV-VLSDPCKEGKFGCSLDIVIHMTAQTS 230
Db 128 -----KRFGLGSSGAVYATVAANAFCGLELSTDERFLAMLATALD- 171
Qy 231 CLAOGKVGSGFVSCAVYGSQRYVRFSPVLISFAQVAVTGLPLNEVITIGKWKDNRKT 290
Db 172 -----PKSGGDLAASWTGG--WIAVQAPDRAF-----VLDAIRRVGVDRITKAPPGASHV 220
Qy 291 EFSIPLMLNLF-----GEPGSGSSTPSMGAVKAKKMWSDPEKAREWQNLSDANLELE 345
Db 221 R-RLPAPRGILLEVGMTEP---ASTASLVS-----DLHRRTRGSASHQREVE 265
Qy 346 TKLNDLSKLADHDVYLRVILKSCSVLTSEKMWLHATEPINEAIKTELLEAREAMLRIRI 405
Db 266 T-----TTDCVRSATVALESGLD--DTSILHETRRARQOLAR--- 299
Qy 406 LMRONGEASVPIEPESQTLDDSTMSAEVLLAGVP-GAGGFDAIFAITLGDSCGTLT- 463

Db 300 ----LDDEVGLIEFTPKLTALD---AAEAVGAAKPSGAGCGIGIALDDEASRDITH 352
Qy 464 --QAMSSHNVAL 474
Db 353 VRQWETAGVLP 365

RESULT 9

Q92FUI PRELIMINARY; PRT; 360 AA.
AC Q92FUI;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Hypothetical protein lin0012.
GN LIN0012.
OS Listeria innocua.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Listeriaceae; Listeria.
OX NCBI_TaxID=1642;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CLIP 11262 / SEROVAR 6A;
RX PubMed=11679669;
RA Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
RA Baquero P., Berche P., Bloeker H., Brandt P., Chakraborty T.,
RA Charbit A., Chetouani F., Couve E., de Darvar A., Dehoux P.,
RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
RA Entian K.-D., Fsihl H., Garcia-del Portillo F., Garrido P.,
RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Haut J., Jackson D.,
RA Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kurapkak G.,
RA Madueno E., Maitounam A., Mata Vicente J., Ng E., Nedjari H.,
RA Nordsiek G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
RA Remmel B., Rose M., Schlueter T., Simoes N., Tierrez A.,
RA Vazquez-Boland J.-A., Voss H., Weiland J., Cossart P.;
RT "Comparative genomics of Listeria species."
RL Science 294:849-852(2001).
DR EMBL; AL596163; CAC95245.1; .
DR ListList; LIN00012; .
DR InterPro: IPR001745; GMPkinase_ATP.
DR Pfam: PF00288; GMP_kinases; 1.
DR TIGRFAMS; TIGR01220; Pmev_kin_Gr_pos; 1.
KW Hypothetical protein; Complete proteome.

Query Match 5.98; Score 151.5; DB 16; Length 360;

Best Local Similarity 20.78; Pred. No. 0.0021;
Matches 100; Conservative 60; Mismatches 145; Indels 177; Gaps 20;

Qy 8 PGKVLMTGGLVLEKPNAGLVLTSTNARFYAIK-----PINEYKP 48
Db 14 PKGLVAGEFYAVESGHAITLAVN-RITTLTLESENEMLIPHYENPVMPIGELKP 72
Qy 49 ESMANKWTDVKTSPOLSRSMYKLSLNHLTLQSVASDSRNPVEHAIOYAIANAHLAT 108
Db 73 DGEHWTFT-----AEAINIAT 88
Qy 109 EKDKESLKLKLLQGLDITILGSDNFYSRNOIESAGLPLTPESLGITLAPFASITFNMAES 168
Db 89 -----TFLKSGIGELTPVKM-----VITETLIDQ 112
Qy 169 NGANSKPEVAKTGIGSSAAMTTAVVAALL--HYLGVNVLSDPCKEGKFGCSLDIVIHMTA 226
Db 113 SG-----AKYGSSAATYAVAINALMTKFPISMLK-----KLA 150
Qy 227 QTSKCLAOGKVGSGFVSCAVYGSQRYVRFSPVLISFAQVAVTGLPLNEVITIGKWKD 286
Db 151 ALSHLYVVGSCSGDIAACMYGWIATYTFQDEWKH-RLAYKSID-----WFKKEPMP 203
Qy 287 NKRTFESLPLMLNLFLEPGSGSSTPSMGAVKAKKMWSDPEKAREWQNLSDANLELET 346
Db 204 MLIETLEPEVPPTFSVGTGT-PVSTGLVSGIHAFKQED--SKNYOHFL-----T 251

QY 347 KLNDLSKLANDHMDVYLRYVTKSCSVLTSEKWLHATEPINEALIKELLEAREMLRIRIL 406
 Db 252 RNEIMK-----QIIQAFH--TKDELLYSSIKENRRILOEL----- 286
 QY 407 MRQGEASVPIEPESOTOLDSTMSAEVLLAG-VPGAGGPAIFAIT-LGDSGRTLQ 464
 Db 287 -----GTAGVNIETSLKELAD--SAENMGAGKSSGSGGCGIAFSKTELAEKLVN 339
 QY 465 AW 466
 Db 340 EW 341

RESULT 10

Q8YAV1 PRELIMINARY; PRT; 359 AA.
 ID O8YAV1
 AC O8YAV1;
 DT 01-MAR-2002 (TReMBLrel. 20, Created)
 DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)
 DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
 DE Hypothetical protein lmo0012.
 GN LMO0012.
 OS Listeria monocytogenes.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
 OC Listeriaceae; Listeria.
 OX NCBI_TaxID=1639;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=EGD-E / SEROVAR 1/2A;
 RX MEDLINE=21537279; PubMed=11679669;
 RA Glaser P., Frangeul L., Buchrieser C., Rusnock C., Amend A.,
 Baquero F., Berche P., Blocher H., Brandt P., Chakraborty T.,
 Charbit A., Chetouani F., Couve E., de Daruvar A., Dehoux P.,
 Dommann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
 Ertlin K.-D., Esili H., Garcia-del Portillo F., Garrido P.,
 Gautier L., Goebel W., Gomez-Lopez N., Hain T., Haut J., Jackson D.,
 Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kurapat G.,
 Madueno E., Maitounam A., Mata Vicente J., Ng E., Nedjari H.,
 Nordiel G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
 Remmel B., Rose M., Schlueter T., Simoes N., Tierrez A.,
 RA Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.;
 RT "Comparative genomics of Listeria species."; Science 294:849-852(2001).
 RL Science 294:849-852(2001).
 DR EMBL: AL591973; CAC98227.1;
 DR Listlist; LMO0012;
 DR InterPro: IPR001745; GHMPkinase_ATP.
 DR Pfam: PF00288; GHMP_kinases; 1.
 DR TIGRPFAM: TIGR01220; Pmev_kin_Gr_pos; 1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 359 AA; 39981 MW; 15156C40BD997A1C CRC64;

Query Match 5.8%; Score 149.5; DB 16; Length 359;
 Best Local Similarity 20.5%; Pred. No. 0.003;
 Matches 99; Conservative 60; Mismatches 146; Indels 177; Gaps 20;

QY 8 PGVLMGTGGLVLEKPNAGLVLTNARFYIVK-----PINEVKP 48
 Db 13 PGKLYVAGEAYVSSGHTALTAVN-RYITLTLEDSEKRNELMIPHYENPVSWPGELKP 71
 QY 49 ESNAWKMTDVKLTSPQSRSMYKLSLNHLTLQSVASDSRNPVEHAIOYATAAALAT 108
 Db 72 DGEHWFT-----AEAINIAT 87
 QY 109 EKDKESLHKLLOGLDITILGSDNDFYSRNOIESAGLPLTPESIGTLAPASITFNAES 168
 Db 88 -----TFKSEGIELPYVM-----IITELIDQ 111
 QY 169 NGANSKPEVAKTGLGSSAATTAVALL--HYLGVDLSDPCKEGKFGCSDDLVIHMA 226
 Db 112 SG-----AKYGLGSSAATAVAINALMTKRYPEISMLK-----KFA-----KLA 149
 QY 227 QTHCLAQGVGSGFVSCAVVYSQRVRSPEVLSPFAQVAVYGLPLANEVIGTILKKKWD 286

Db 150 ALSHLVYQNGSCGDIASCMYGIWATTPDQEVKH-RLAYSLF-----WFMKEPWP 202
 QY 287 NKRTFELPLMLNLFLEPSSGSSTPSNYGAVKMQMSPDEKARENQMUNISANILET 346
 Db 203 MLDIETLEEFVPPFVSGWGT-PVSTCKLYSQIHAFQED-----SKNYQHL-----T 250
 QY 347 KLNDLSKLANDHMDVYLRYVTKSCSVLTSEKWLHATEPINEALIKELLEAREMLRIRIL 406
 Db 251 RNEIMK-----QIIQAFH--TKDELLYSAIKENRRILOEL----- 285
 QY 407 MRQGEASVPIEPESOTOLDSTMSAEVLLAG-VPGAGGPAIFAIT-LGDSGRTLQ 464
 Db 286 -----GTAGVNIETSLKELAD--SAENMGAGKSSGSGGCGIAFSKTELAEKLVN 338
 QY 465 AW 466
 Db 339 EW 340

RESULT 11

Q9FD62 PRELIMINARY; PRT; 361 AA.
 ID Q9FD62
 AC Q9FD62;
 DT 01-MAR-2001 (TReMBLrel. 16, Created)
 DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
 DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
 DE Phosphomevalonate kinase.
 GN MVAK2.
 OS Enterococcus faecium (Streptococcus faecium).
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
 OC Enterococcaceae; Enterococcus.
 OX NCBI_TaxID=1352;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20353468; PubMed=10894743;
 RA Wilding E.I., Brown J.R., Bryant A.P., Chalker A.F., Holmes D.J.,
 Ingraham K.A., Iordanesco S., So C.Y., Rosenberg M., Gwyn M.N.;
 RT "Identification, evolution and essentiality of the mevalonate pathway
 for isopentenyl diphosphate biosynthesis in gram-positive cocci";
 RL J. Bacteriol. 182:4319-4327(2000).
 DR EMBL: AF290095; ANG02447.1;
 DR InterPro: IPR000515; BPD_transp.
 DR InterPro: IPR001745; GHMPkinase_ATP.
 DR InterPro: IPR001459; Mev_gal_kin.
 DR Pfam: PF00288; GHMP_kinases; 1.
 DR PRINTS: PRO0959; MEVGALKINASE.
 DR TIGRPFAM: TIGR01220; Pmev_kin_Gr_pos; 1.
 DR PROSITE: PS00402; BPD_TRANSF_INN_MEMBER; UNKNOWN.1.
 KW kinase.
 SQ SEQUENCE 361 AA; 39573 MW; 7802796FC68A0B81 CRC64;

Query Match 5.7%; Score 147.5; DB 2; Length 361;
 Best Local Similarity 21.2%; Pred. No. 0.0044;
 Matches 102; Conservative 64; Mismatches 171; Indels 145; Gaps 19;

QY 3 VVASAPGVLMTGGYLVLEKPNAGLVLTNARFYIVKPINEVKPESNAWKMTDVKLT 62
 Db 2 IEVASAPGKLYIAGEYAVETGHPAVIAAVDQFTVTVESARKVGSIOAQVSCMPVWMT 60
 QY 63 POLSRRESMYKLSLNHLTLQSVASDSRNPVEHAIOYATAAALATEKQESLHKLLOG 122
 Db 61 ---RRNGELVLDLR-----ENPF-----HTLAIARL-TEKYAQEKNIL-- 96
 QY 123 LDITILGSDNDFYSRNOIESAGLPLTPESIGTLAPASITFNAESNGANSKPEVAKTGL 182
 Db 97 -----SFYDLK-----VISELSSNG-----KRYGL 117
 QY 183 GSSAAMTAVVAALHLHYLGVDLSDPCKEGKFGCSDDLVIHMAQTSCLAQGVSGGFD 242
 Db 118 GSCAVATVATKALNVFYAL-----MLSOLEIFKIALAN--LAVQDNCSGCD 163
 QY 243 VSCAVVYSQRVRSPEVLSPFAQVAVYGLPLANEVIGTILKKKNDKRTESPLPL----- 287

Db 164 IAAACYGG--WIAFSPEDHMLQEOETQHSISELLADWPG-----LSIEPLIAPD 213
QY 298 MNLELGRSGSGSTPSMGAVKKWQMSDEPKARENNQNSDANLELETLNDLSKLAKD 357
Db 214 LRLIGTGS-PASTSDLYOV-----HRSRED----- 240
QY 358 HMDVYLKRVKSCSVLTSEKWLHATEPINEAI--IKE--LLEAREAMLRIIRLROMGEA 413
Db 241 -----KMAVAQFLKKNSTECVNMINGFKENNNTLLIQMRKRLDLSAI 288
QY 414 ASVPIDEPESOTQLLDSTMSAEVLLAGVPGAGFDALFATLGDG--TKLTQANSSHNL 472
Db 289 TGVVIEFPALNKLCLNAEOYEGA--AKSGAGGDCGIVDQKSGILPMSAMEKAET 346
QY 473 AL 474
Db 347 PL 348

RESULT 12

Q9FD72 PRELIMINARY; PRT; 358 AA.
AC Q9FD72;
DT 01-MAR-2001 (TREMblrel. 16, Created)
DT 01-MAR-2001 (TREMblrel. 16, last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, last annotation update)
DE Phosphomevalonate kinase.
GN MYAK2.
OS Staphylococcus epidermidis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=1282;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20353468; PubMed=10894743;
RA Wilding E.I., Brown J.R., Bryant A.P., Chalker A.F., Holmes D.J.,
RA Ingraham K.A., Iordanescu S., So C.Y., Rosenberg M., Gwynn M.N.;
RT "Identification, evolution and essentiality of the mevalonate pathway
for isopentenyl diphosphate biosynthesis in gram-positive cocci".
RL J. Bacteriol. 182:4319-4327(2000).
DR EMBL: AF290091; AAC02437.1;
DR InterPro: IPR001745; GHMPkinse_ATP.
DR InterPro: IPR001459; Mev_gal_kin.
DR Pfam: PF00288; GHMP_kinases; 1.
DR PRINTS: PR00959; MEVGALKINASE.
DR TIGRFRAMS: TIGR01220; Pmev_kin_Gr_pos; 1.
KW kinase.
SQ SEQUENCE 358 AA; 39886 MW; 608EBDADF5774EC2 CRC64;

Query Match 5.5%; Score 142.5; DB 2; Length 358;
Best Local Similarity 19.8%; Pred. No. 0.011;
Matches 94; Conservative 72; Mismatches 167; Indels 141; Gaps 20;

QY 3 VVASAPGKVLMTGGYLVLEKPNAGVLSTNARFYAIVKPIINEEVKPSMAWKTDVKLTS 62
Db 2 IGVKAPGKLYIAGYVATEPGYSILIAVNRFTVATIEASNK-----VEGSIHS 50
QY 63 POLSRESM-KKLSLNHLTLOSASDSRNPFVEHAIOYALIAAHLATEKDESLHKLLQ 121
Db 51 KTLATPEVAKFDRNEDRLDISDYQAAKO-----LKYVTAIEV----- 87
QY 122 GLDITILGSDNFYSYRNQIESAGLPLTPESLGTLPAPASITFNAESNGANSKPEVAKTG 181
Db 88 -----FEDYVYSCNNMLKHFL-----TDSNLADNSG-----QKYG 119
QY 182 LGSSAAMTAVVAALLHYLVNDLSDPCKEKGCGSDLDVYHMTAQTSHCLAQKVGSGF 241
Db 120 LGSSAAVLAVYVVALNEFYG-LELSNL-----YIKLAVIANMKLQSLSSCG- 165
QY 242 DVSCAVY-GSORVYRSPPEVLSPFAOVAVTGLPLNEVIGTILKGMKDKRTE-FSLPLLMN 299
Db 166 DIAVSIVSGMLAVSTFDHDMVK-QQMEETS-----VNDVLEKMPGLHIEPLQAPENME 218

QY 300 LFLGEPGSGGSTPSMGAVKKWQMSDEPKARENNQNSDANLELETLNDLSKLAKDH 359
Db 219 VLIQWTSPPASS-PHYSEVKRLK-SDP-----SFY 247
QY 360 DVLKRVKSCSVLTSEKWLHATEPINEAIKELLEAREAMLRI-RILRQMGASAVPI 418
Db 248 GDFLDQSHAC-----VESLQAFKTNNIKGVQKIRIRRIILOSMDNEASVEI 295
QY 419 EPESOTQLLDSTMSAEVLLAGVPGAGFDALFATLGDGSKTLQAMSSHNL 472
Db 296 EFDKLRKLDVGEKHGA--SKTSGAG-----GDCGITTINKVYIDRNT 338

RESULT 13

Q9FD83 PRELIMINARY; PRT; 358 AA.
AC Q9FD83;
DT 01-MAR-2001 (TREMblrel. 16, Created)
DT 01-MAR-2001 (TREMblrel. 16, last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, last annotation update)
DE Phosphomevalonate kinase.
GN MYAK2.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20353468; PubMed=10894743;
RA Wilding E.I., Brown J.R., Bryant A.P., Chalker A.F., Holmes D.J.,
RA Ingraham K.A., Iordanescu S., So C.Y., Rosenberg M., Gwynn M.N.;
RT "Identification, evolution and essentiality of the mevalonate pathway
for isopentenyl diphosphate biosynthesis in gram-positive cocci".
RL J. Bacteriol. 182:4319-4327(2000).
DR EMBL: AF290087; AAC02426.1;
DR InterPro: IPR001745; GHMPkinse_ATP.
DR InterPro: IPR001459; Mev_gal_kin.
DR Pfam: PF00288; GHMP_kinases; 1.
DR PRINTS: PR00959; MEVGALKINASE.
DR TIGRFRAMS: TIGR01220; Pmev_kin_Gr_pos; 1.
KW kinase.
SQ SEQUENCE 358 AA; 40203 MW; 67EF432FA155FEF5 CRC64;

Query Match 5.3%; Score 136.5; DB 2; Length 358;
Best Local Similarity 21.5%; Pred. No. 0.03;
Matches 103; Conservative 60; Mismatches 180; Indels 135; Gaps 23;

QY 3 VVASAPGKVLMTGGYLVLEKPNAGVLSTNARFYAIVKPIINEEVKPSMAWKTDVKLTS 62
Db 2 IGVKAPGKLYIAGYVATEPGYSILIALDRF--VTATIEEARQ-----YKGT----- 47
QY 63 POLSRESMKKLSLNHLTLOSASASDS---RNPFEHAIQYALIAAHLATEKDESLHKLL 119
Db 48 -----HSAKALHNHPVTSRDEDSIVISPHAKOKLNYVTAIEFEO----- 90
QY 120 LOGDITILGSDNFYSYRNQIESAGLPLTPESLGTLPAPASITFNAESNGANSKPEVAK 179
Db 91 -----YAKSCDLAMKHFLITIDS-----NDDDSNG-----HK 117
QY 180 TGLSSAAMTAVVAALLHYLVNDLSDPCKEKGCGSDLDVYHMTAQTSHCLAQKVGSG 239
Db 118 YGLSSAAVLAVYVVALNEFYDM-----KLSNL-YIKLAVIANMKLQSLSSC 164
QY 240 GFDVSCAVY-GSORVYRSPPEVLSPFAOVAVTGLPLNEVIGTILKGMKDKRTE-FSLPL 297
Db 165 G-DIAVSIVSGMLAVSTFDHDMVKH-QIEDT-----TVEVVLKMPGLHIEPLQAPEN 216
QY 298 MNLELGRSGGSTPSMGAVKKWQMSDEPKARENNQNSDANLELETLNDLSKLAKD 357
Db 217 MEVLIGMTGSPASS-PHYSEVKRLK-SDP-----S 245
QY 358 HMDVYLKRVKSCSVLTSEKWLHATEPINEAIKELLEAREAMLRIIRLROMGEASAVP 417

Db 246 FYGDFLEDSHRC-----VEK-LIHAFKTNINIKGVQKVRON-----RTTIQRMDKEATYD 294
 QY 418 IEPESQTOILDSTMSAGVLLAGVPGAFDAIFATITLGD-SGTKLTQAWSSHVAL 474
 Db 295 IETEKILKYLCDIAEKYHGA--SKTSAGGSGDCITTIINKNVDKYEKIDENKTKIKRPL 350

RESULT 14

Q9KMF7 PRELIMINARY; PRT; 360 AA.

AC 09KMF7;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Phosphomevalonate kinase.
 GN PMEVK.
 OS Kitiatsospora griseola (Streptomyces griseoloporeus).
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycinae; Streptomycetaceae; Kitiatsospora.
 OX NCBI_Taxid=2064;
 RN (1)
 RP SEQUENCE FROM N.A.
 RA Dairi T.;
 RT "Cloning of a gene cluster encoding enzymes responsible for the
 RT mevalonate pathway from a terpenoid antibiotic-producing Streptomyces
 RT strain."
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB037907; BAB07819.1;
 DR InterPro: IPR001745; GHMPkinase.ATP.
 DR InterPro: IPR001459; Mey_gal_kin.
 DR Pfam: PF00288; GHMP_kinases.1.
 DR PRINTS: PR00959; MEVGALKINASE.
 DR TRIPFAMS: TIGR01220; pmev_kin_gr_pos; 1.
 KW KINASE.
 SQ SEQUENCE 360 AA; 37833 MW; FD194D64355C7ZDE CRC64;

Query Match 5.3%; Score 136.5; DB 2; Length 360;

Best Local Similarity 19.5%; Pred. No. 0.031; Mismatches 174; Indels 159; Gaps 17;

Matches 97; Conservative 67; Mismatches 174; Indels 159; Gaps 17;
 QY 2 AVVASAPGVLTGTYLVEKPNAGVLSTNARFYAIVK-----INEVVPESMAWK 54
 Db 6 AVTRARPKLFAVEAYVPEGRALIVNDRVYTVTSNGARPVYSSDIDAGVYHNP 65
 QY 55 WTDVKTSPQSLRESMYKLSLNLTLQSVASDSRNPVEHAIOYAIAMHATEKDKES 114
 Db 66 WODGRLT-----GGTGPHVVAAYE--TVARLAEK----- 94
 QY 115 LKHLLOGDITTLGSDNFYSYRNQIESAGLPLTPESLGTLPFASTTFNAESNGANSK 174
 Db 95 -----GNSVPLG--WSISSTLHEDG----- 113
 QY 175 PEVAKTGLSSAAMTAAVVAALHYLVGVNLDSPCKEKGCSDDLVIHMIQTSCHLAQ 234
 Db 114 ---RKGLGSSGAVYATVSAVAHAGL---ELTADEER-----RTALIAS 153
 QY 235 GKV---GSGFDVSCAVYSGORVYRSPSEVLSFAQVAVTGLPLNEVIGTTLKGWKNRTE 291
 Db 154 ARIDPREGSGDITFTWGWGIAVYR-APD-----RDVILDLTRQGVDEALRAPMPEFSVR 207
 QY 292 FSLPPLMNLFLGEPGSGSSTPSMVGAVKKMKMSDEPKARENNQNSDANLELETNLNDL 351
 Db 208 LSPRNLCLEVGVGTGNPVSTT-----SLLTLD 234
 QY 352 SKLAKDMDVYLVIRKSCVLTSEKWLHATEPINEAIKELLEAREAMLRIILKROMG 411
 Db 235 HRTWNGSAFYRRYVATGELV-DAAVIALEDDTEGLLRQVRAHENVYR---LDDDEVG 290
 QY 412 EAASVPIEPESQTOILDSTMSAGVLLAGVPGAFDAIFATITLGDGSK-----LTQAW 466
 Db 291 LGIFTE-----ELFALCAIAERAGAAKRSAGGDC--GIALDDAEARVDRSPILHRQW 341
 QY 467 SSHNVALLVREDPHGV 483

Db 342 AAAGVLLPLVSPATEGV 358

RESULT 15

Q99W20 PRELIMINARY; PRT; 358 AA.

AC 099W20;
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Phosphomevalonate kinase.
 GN WVAK2 OR SAV0392 OR SA0549.
 OS Staphylococcus aureus (strain Mu50 / ATCC 700699), and
 OC Staphylococcus aureus (strain N315).
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
 OC Staphylococcus.
 OX NCBI_Taxid=158878, 158879;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC SPECIES-S.aureus (strain Mu50), and S.aureus (strain N315);
 RX MEDLINE-21311952; PubMed-11418146;
 RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
 RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Iian J.-Q., Ito T.,
 RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,
 RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
 RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
 RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
 RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
 RT "Whole genome sequencing of methicillin-resistant Staphylococcus
 RT aureus."
 RL Lancet 357:1225-1240(2001).
 DR EMBL; AP003359; BAB56754.1;
 DR EMBL; AP003131; BAB41781.1;
 DR InterPro: IPR001745; GHMPkinase.ATP.
 DR InterPro: IPR001459; Mey_gal_kin.
 DR Pfam: PF00288; GHMP_kinases.1.
 DR PRINTS: PR00959; MEVGALKINASE.
 DR TRIPFAMS: TIGR01220; pmev_kin_gr_pos; 1.
 KW KINASE; Complete proteome.
 SQ SEQUENCE 358 AA; 40217 MW; 19071320FA51AFEL CRC64;

Query Match 5.2%; Score 134.5; DB 16; Length 358;

Best Local Similarity 21.8%; Pred. No. 0.043; Mismatches 181; Indels 133; Gaps 23;

Matches 104; Conservative 59; Mismatches 181; Indels 133; Gaps 23;

QY 3 VVASAPGVLTGTYLVEKPNAGVLSTNARFYAIVK-----INEVVPESMAWKTDVTKLS 62
 Db 2 IOVKAPGKLYIAGEVAVTE-PCYKSVLALDR-----VT 35
 QY 63 POLSRESMYKLSLNLTLQSVASDSRNPVEHAIOYAIAMHATEKDKESLHKLLOG 122
 Db 36 ATTEADQYKGTHTSKALHNHPYTFSRD--EDSI--VISDPHAAKQ----- 77
 QY 123 LDTTLGSDNFYSYRNQIESA--GLPLTPESLGTLPFASTTFNAESNGANSKPREVAKT 180
 Db 78 LNVVTAIEIFEOYANSCDIAMKHFLLTDS-----NLDDSGN-----HXY 118
 QY 181 GLGSSAAMTAAVVAALHYLVGVNLDSPCKEKGCSDDLVIHMIQTSCHLAQKVGSG 240
 Db 119 GLGSSAALVYAVYKLVNEYTD-----KLSNL-YIYKLVANMKLQSLSSSG 165
 QY 241 FDVSCAVY-GSORVYRSPSEVLSFAQVAVTGLPLNEVIGTTLKGWKNRTE-FSLPPLM 298
 Db 166 -DIAYSVYSGMLAYSTFDHEWVKH-QIEDT-----TVEEVILKNPGLHIEPLAQPEMN 217
 QY 299 NLFLEPGSGSSTPSMVGAVKKMKMSDEPKARENNQNSDANLELETNLNDLSKLAKH 358
 Db 218 EYLIGWTGSPASS-PHFVSEVRLK-SDP-----SF 246
 QY 359 MDVYLVIRKSCVLTSEKWLHATEPINEAIKELLEAREAMLRIILKROMGEASVPI 418
 Db 247 YGDFLEDSHRC-----VEK-LIHAFKTNINIKGVQKVRON-----RTTIQRMDKEATYDI 295

Wed, May 07 14:13:43 2003

us-09-988-863a-2.rsp

Page 9

QY 419 EPESQTDLDSTMSAEVLLAGVPGAGCFDAIFATLLGD-SCTKLTQANSSHNLAL 474
| | | : ||| | | : | : | : |
Db 296 ETEKLYLCDIAEKYHGA--SKTSGAGGDCGTTINKDVREKIIYDEMTKHGIKPL 350

Search completed: April 26, 2003, 12:57:02
Job time : 92 secs

THIS PAGE BLANK (001910)

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using SW model

Run on: April 26, 2003, 12:54:00 ; Search time 22 Seconds
(without alignments)
2206.722 Million cell updates/sec

Title: US-09-988-863A-2
Perfect score: 2581
Sequence: 1 MAVVASAPGKVLMTGTYLV.....ESGDPRTTCITSGVSITLH 505

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: PIR_73:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2581	100.0	505	2 C86443	unknown protein [i
2	493	19.1	426	2 T38650	probable phosphome
3	453.5	17.6	451	2 S57588	phosphomevalonate
4	230	8.9	323	2 P90479	(Phospho) mevalona
5	151.5	5.9	360	2 A61434	mevalonate kinases
6	149.5	5.8	359	2 A61806	mevalonate kinases
7	134.5	5.2	358	2 B89828	phosphomevalonate
8	130	5.0	816	2 T21713	hypothetical prote
9	124.5	4.8	317	2 F70185	phosphomevalonate
10	123	4.8	378	1 S42088	mevalonate kinase
11	117.5	4.6	721	2 A12447	hypothetical prote
12	117	4.5	396	1 A42919	mevalonate kinase
13	117	4.5	557	4 S42226	hypothetical large
14	116	4.5	335	2 A71042	probable mevalonat
15	115	4.5	816	2 B86705	ATP-dependent prot
16	112.5	4.4	478	2 D97447	hypothetical prote
17	112.5	4.4	478	2 A62665	FAD dependent oxid
18	112.5	4.4	902	2 G83635	probable ClpA/B-ty
19	112	4.3	287	1 DKRTBH	libeta-hydroxyster
20	112	4.3	335	2 H75172	mevalonate kinase
21	111.5	4.3	395	1 A35629	mevalonate kinase
22	110.5	4.3	1417	2 H83132	probable sensor/re
23	109.5	4.2	292	1 I56604	libeta-hydroxyster
24	109	4.2	662	2 S58298	APPase - pepper (f
25	108.5	4.2	615	2 E98310	rhizobial secret
26	108.5	4.2	615	2 AF2972	rhizobial secret
27	108.5	4.2	659	2 AB0114	arginine decarboxy
28	108.5	4.2	1477	2 B43855	high-molecular-wet
29	107	4.1	463	2 T36810	probable integral

30	107	4.1	1442	2 S57160	sulfite reductase
31	106	4.1	698	1 S31714	MDP-protein ADP-r
32	106	4.1	1087	2 D84767	probable pre-mRNA
33	106	4.1	1436	2 S67655	probable membrane
34	105.5	4.1	1341	2 T17285	hypothetical prote
35	104.5	4.0	397	2 F84345	hypothetical prote
36	104.5	4.0	1983	2 AC1922	two-component hybr
37	104	4.0	574	2 T51799	libeta-hydroxyster
38	103.5	4.0	284	2 A69536	mevalonate kinase
39	103.5	4.0	291	1 A55573	mevalonate kinase
40	103	4.0	710	2 S28014	oud protein - Erw
41	103	4.0	788	2 B84857	hypothetical prote
42	102	4.0	764	2 H82737	malate oxidoreduct
43	102	4.0	812	2 T19446	hypothetical prote
44	101.5	3.9	662	2 T17339	hypothetical prote
45	101	3.9	617	2 AD3123	conserved hypotnet

ALIGNMENTS

RESULT 1

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001

C:Accession: C86443

R:Rheologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon

Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar,

ansen, N.F.; Hughes, B.; Hutzar, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Malli, R.; Marzia

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719; PMID:11130712

A:Accession: C86443

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-505 <STO>

A:Cross-references: GB:A8005172; NID:911136726; PIDN:AAG31307.1; GSPDB:GN00141

C:Genetics:

A:Map position: 1

Query Match	100.0%;	Score 2581;	DB 2;	Length 505;
Best Local Similarity	100.0%;	Pred. No. 3.8e-191;	Mismatches 0;	Indels 0;
Matches 505;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MAVVASAPGKVLMTGGYLVLEKPNAGLVSTNAREYATVKKPINEBPKESNWKMTDVKL	60	
Db	1	MAVVASAPGKVLMTGGYLVLEKPNAGLVSTNAREYATVKKPINEBPKESNWKMTDVKL	60	
QY	61	TSPQSRSMYKLSLNHLTLOSASDSRNPFEAIOYAIAMAHATEKDESHKLL	120	
Db	61	TSPQSRSMYKLSLNHLTLOSASDSRNPFEAIOYAIAMAHATEKDESHKLL	120	
QY	121	OGDLITILGSNDYFYRNQIESAGLPITPESIGTAPRASITFPNAESNGANSKREYAKT	180	
Db	121	OGDLITILGSNDYFYRNQIESAGLPITPESIGTAPRASITFPNAESNGANSKREYAKT	180	
QY	181	GIGSSAAMTTAVVALLHYLGAVDSDPCBEKFCPSDLDYHMAQSHCLAGKVGSG	240	
Db	181	GIGSSAAMTTAVVALLHYLGAVDSDPCBEKFCPSDLDYHMAQSHCLAGKVGSG	240	
QY	241	FDVSCAVYGSQRYVRFSEVLISFAQVAVTGLPLNEVIGTILGKWNKRTEFSPLPLNL	300	
Db	241	FDVSCAVYGSQRYVRFSEVLISFAQVAVTGLPLNEVIGTILGKWNKRTEFSPLPLNL	300	
QY	301	FLGEGSGSGSSTPSNVGVAKKQMSDPEKARNMNLSDANILETKLNDLSKLAKDMD	360	
Db	301	FLGEGSGSGSSTPSNVGVAKKQMSDPEKARNMNLSDANILETKLNDLSKLAKDMD	360	

QY 361 VLEVRKSCSVLTSEKVNLTATEPINEAIKELLEAREAMLRIILMRONGEASVPIEP 420
 |||||
 Db 361 VLEVRKSCSVLTSEKVNLTATEPINEAIKELLEAREAMLRIILMRONGEASVPIEP 420
 |||||
 QY 421 ESQTOLDSTMSAGVLLAGVPGAGGDAIFAITLGSCKRLTQWAMSHVLLALIVEDP 480
 |||||
 Db 421 ESQTOLDSTMSAGVLLAGVPGAGGDAIFAITLGSCKRLTQWAMSHVLLALIVEDP 480
 |||||
 QY 481 HGVCLSEGDPTTCITSGVSIHLE 505
 |||||
 Db 481 HGVCLSEGDPTTCITSGVSIHLE 505
 |||||

RESULT 2

T38650
 probable phosphomevalonate kinase - fission yeast (Schizosaccharomyces pombe)
 C:Species: Schizosaccharomyces pombe
 C:Date: 03-Dec-1999 #sequence-revision 03-Dec-1999 #text-change 20-Jun-2000
 C:Accession: T38650
 R:Murphy, L.; Harris, D.; Wood, V.; Rajandream, M.A.; Barrell, B.G.
 submitted to the EMBL Data Library, August 1997
 A:Reference number: Z21804
 A:Accession: T38650
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-426 <MUR>
 A:Cross-references: EMBL:AL109739; PIDN:CA852264.1; GSPDB:GN00066; SPDB:SPAC343.01c
 A:Experimental source: strain 972h-; cosmid c343
 C:Genetics:
 A:Gene: SPDB:SPAC343.01c
 A:Map position: 1
 A:Introns: 3/2

Query Match 19.1%; Score 493; DB 2; Length 426;

Best Local Similarity 30.2%; Pred. No. 4.2e-30; Mismatches 170; Indels 100; Gaps 21;

Matches 153; Conservative 83; Mismatches 170; Indels 100; Gaps 21;

QY 1 MAVASAPGKVLMTGVLLEKPNAGLVLTSTNARFYAIKPIINEVVRPESWAMKMTDKL 60
 |||||
 Db 1 MKVTCGAPGKVLINGVILVDPQYSGVIGLTAAGVASTTLTD-----KCGYRV 51
 |||||
 QY 61 TSPQ-LSRESMYKLSLN-----HLTQSVASDSRNPFEVHAIOYATAAHLATEKDE 113
 |||||
 Db 52 KSPQFINAEMLYNIDWTVSPRIHQIYENCLEKNPMPFVLALFYVI-NYFESTGR--- 107
 |||||
 QY 114 SLHLLLOGDITLIGSNDIFYRNQIESAGLPLTPESLGLTAP---FASTFMAESNG 170
 |||||
 Db 108 --QPLCMODLVTLQVDNAY-HQPO-----LKPQO--TSYKFNPLNCTLG----- 149
 |||||
 QY 171 ANSKPEVAKTGIGSSAAMTTAVVALLHYLGVDLSDPC--KEGKFGCSDLDVTHMIAQT 228
 |||||
 Db 150 ----QVHKTGIGSSAAMTISLISL--FLSLRLTDTDTGKSLKIDSTVIYHNLAOI 202
 |||||
 QY 229 SHCIAQGVSGFDVSCAVVGSQRYVRPSEVLSPAQV---AVTGLPLNEVIGTILK 284
 |||||
 Db 203 AHCAGCGVSGFDVSCAVVGSQRYVRPSEVLSPAQV---AVTGLPLNEVIGTILK 284
 |||||
 QY 265 WDNKRTFESLPLNLNLFLEPGSGGSTPSPVAVKWKQMSDPKARENMNQLSDANIEL 344
 |||||
 Db 263 WSDV-VPPQPLATYCLLMGDV-AGGSSTPGVAKVQOMKQNPRESK---NCFD----- 311
 |||||
 QY 345 ETKLNDLSKLAKDHWVLYRV--IKSCSVLTSEKVNLTATEPINEAIKELLEAREAML 402
 |||||
 Db 312 -----DLSRVLSLTKNCL-----SSSLDSELOSQSRISRIILOR 347
 |||||
 QY 403 IRIIMRONGEASVPIEPESOTQLDSTMSAEGVLLAGVPGAGGDAIFAITLGDSC--T 460
 |||||
 Db 348 ITV-----EAKVDIEPLKQTNILNDNEQLPGVIGVPGAGGDAOFCLAINHTTEIE 400
 |||||
 QY 461 KLTQAMSHVLLALIVEDPHGVCL 486
 |||||
 Db 401 NVITKMDGCVVPMQVSPAFDGLAVE 426
 |||||

RESULT 3

S57588
 Phosphomevalonate kinase (EC 2.7.4.2) - Yeast (Saccharomyces cerevisiae)
 N:Alternate names: protein YN959.02; protein YMR220W
 C:Species: Saccharomyces cerevisiae
 C:Date: 19-Oct-1995 #sequence-revision 03-Nov-1995 #text-change 21-Jul-2000
 C:Accession: S57588; A39606
 R:Skellon, J.; Churcher, C.M.
 submitted to the EMBL Data Library, June 1995
 A:Reference number: S57587
 A:Accession: S57588
 A:Molecule type: DNA
 A:Residues: 1-451 <SKE>
 A:Cross-references: EMBL:Z49939; NID:9887599; PID:9887601; MIPS:YMR220W
 A:Experimental source: strain AB972
 R:Teay, Y.H.; Robinson, G.W.
 Mol. Cell. Biol. 11, 620-631, 1991
 A:Title: Cloning and characterization of ERG8, an essential gene of Saccharomyces cer
 A:Reference number: A39606; MUID:91117228; PMID:1846667
 A:Accession: A39606
 A:Molecule type: DNA
 A:Residues: 1-212 'R', 214-417 'PLMF', 422 'D', 424 <TSA>
 A:Cross-references: EMBL:M63648; NID:9553127; PIDN:AAA34596.1; PID:9171479
 C:Genetics:
 A:Gene: SGD:ERG8
 A:Cross-references: SGD:S0004833; MIPS:YMR220W
 A:Map position: 13R
 C:Keywords: phosphotransferase

Query Match 17.6%; Score 453.5; DB 2; Length 451;

Best Local Similarity 30.1%; Pred. No. 5.1e-27;

Matches 155; Conservative 72; Mismatches 185; Indels 103; Gaps 18;

QY 6 SAPKVLMTGGLVLEKPNAGLVLTSTNARFYAIKPIINE-----EVKESWAMK---W 55
 |||||
 Db 8 SAPKALLAGVYLVDKRYEAFVGLSARMAVAHPGSLQSDKFEVRKSKQKGEW 67
 |||||
 QY 56 TDVLTSPQLSRESMYKLSLNHLTLOS-----VSASDSRNPFEVHAIOYATAAHLATEKD 111
 |||||
 Db 68 -----LKHISPKGFIPVSGSKNPFIEVIANVSYF----- 101
 |||||
 QY 112 KESLKLLOGDITLIGSNDIFYRNQIESAGLPLTPESLGLTAPASTFMAESNGA 171
 |||||
 Db 102 KPNMDYCNRLFEYIDIFSD--AYHSQEDS-----VTEHRG-----NRLSFS 144
 |||||
 QY 172 NSKPEVAKTGIGSSAAMTTAVVALLHYLGVDLSDPCKEGKRECCSDLDVTHMIAQTS 231
 |||||
 Db 145 HRIEVPKRTGIGSSAGLVTLTALASF--VSDLENNVDKIR-----EYIHNLAQVAHC 197
 |||||
 QY 232 LAQKVGSGFDVSCAVVGSQRYVRPSEVLSPAQVAVTGLPLNEVIGTILK----- 284
 |||||
 Db 198 OAGKIGSGFDVSCAVVGSQRYVRPSEVLSPAQVAVTGLPLNEVIGTILK----- 284
 |||||
 QY 285 --WDNRTFESLPLNLNLFLEPGSGGSTPSPVAVKWKQMSDPKARENMNQLSDAN 341
 |||||
 Db 248 EEDWNITITKSNHLPSGLTLMWMDI-KNGSETVKLVQYKKNWYDSHMPESLKIYTELDHAN 306
 |||||
 QY 342 LELETKLNDLSKLAKDHWVLYRVKSC--SVLTSEKVNLTATEPINEAIKELLEARE 399
 |||||
 Db 307 SRFDGLSKLDRIHETHDYSDQIFESLERDCTQRY-----PEITTEVDA 353
 |||||
 QY 400 MLRIIMRONGEASVPIEPESOTQLDSTMSAEGVLLAGVPGAGGDAIFAITLGDSC 459
 |||||
 Db 354 VATIRSFRTKTESGADIEPPVOTSLDDCQITLKVLTCLIPAGAGYDAIATVTKQD 413
 |||||
 QY 460 TKLTQAMSSH--NVLLALIVEDPHGVCL 492
 |||||
 Db 414 LRAQTANDKRFKSVQWMDVTLQADWGVKREK-DPEI 447
 |||||

RESULT 4

F90479
(Phospho) mevalonate kinase, probable [Imported] - Sulfolobus solfataricus
C:Species: Sulfolobus solfataricus
C:Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 24-May-2001
C:Accession: F90479
R:She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Aayez, M.J.; Chan-
Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, R.
arrett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.
Submitted to GenBank, April 2001
A:Description: Sulfolobus solfataricus complete genome.
A:Reference number: A99139
A:Accession: F90479
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-323 <KUR>
A:Cross-references: GB:AE006641; NID:g13816374; PIDN:AAK43093.1; GSPDB:GN00155
C:Genetics:
A:Gene: SS02988

Query Match 8.9%; Score 230; DB 2; Length 323;
Best Local Similarity 22.2%; Pred. No. 5,3e-10;
Matches 110; Conservative 72; Mismatches 129; Indels 184; Gaps 21;

QY 3 VVASPKVLTGGLVLEKPNAGL--VLSTNAPFAIVKPIINEVKEPESWAKMTDVKL 60
D 2 IKVASPKKILMIGSYV---FGISHYIAVNRKVSCLREIK----- 42
QY 61 TSPQSESM-YKSLNHLITQSVASDSRNPFEHAIQYIAAHATEKDEKSLHL 119
D 43 -----KSLLEFHTSYGF-----KNSGELINSVL-----DFFRRLSO-L 77
QY 120 LOGIDITILGSNDPFSYRNOIESAGLPLTPESLGLTAPFASITFNAESNGANSKPEVAK 179
D 78 PQGEIDLNDKEELI----- 97
QY 180 TGLGSSAAMTAVVAALLHYLVLDSPCKEGFGCSLDLVIMIAQTSICLAQKVG 239
D 98 TGLGSSAAMTAVVAALLHYLVLDSPCKEGFGCSLDLVIMIAQTSICLAQKVG 239
QY 240 GFDVSCAVGSGQRYRFSPEVLSFAQVAVTGLPLNEVIGTILKKMKRFEFLPL-- 297
D 145 GFDIASAVFGSIYKRT-----DLDKMPFEKLN 176
QY 298 --MNLFLGEPGSGSSTPSMGAVKKM-QMSDEPKAREMNONLSDANLEETKINDSKL 354
D 177 GNYMMLGFTGKSET---VGLVRKFEKSNLDDFKEMILDEBNV-MAIKLIKLNKL 231
QY 355 --AKDHMDVYLKVIKSCSVLTSEKVLHATEPINEAIIKELLEAREAMLRIILRMQGE 412
D 232 DEAVEH-----IKLGRKYLVNIAERI--VGKLVSKMEELIKI----- 268
QY 413 AASVPIPEESOTQLDSTMSAGVLAGVPGAGFCAIFATLDSGKLTQAWSHNVL 472
D 269 -----AEEEGALVALSPGAGGSDIFA--LGNDLNRVEAWSKRCIF 308
QY 473 ALVREDPHVGLCES 487
D 309 IIDVKED-EGRLRES 322

RESULT 5
AE1434
mevalonate kinases homolog lln0012 [Imported] - Listeria innocua (strain Clip11262)
C:Species: Listeria innocua
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001
C:Accession: AE1434
R:Glaser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker,
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurgeit, O.; Entian, K.D.; Fsihl, H.;
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Krefit, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maltournam, A.; Ma-
ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlend,
A.; Title: Comparative genomics of Listeria species.

A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AE1434
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-360 <GLA>
A:Cross-references: GB:AL592022; PIDN:CA095245.1; PID:g16412433; GSPDB:GN00178
A:Experimental source: strain Clip11262
C:Genetics:
A:Gene: lln0012

Query Match 5.9%; Score 151.5; DB 2; Length 360;
Best Local Similarity 20.7%; Pred. No. 0.00073;
Matches 100; Conservative 60; Mismatches 145; Indels 177; Gaps 20;

QY 8 PGKVLMTGGLVLEKPNAGLVLSTNAPFAIVK-----PINEVKP 48
D 14 PGKLVAGEVAVESGHAILTAVN-RITVLTLEDSENELMIDHYENPWSMPGELKP 72
QY 49 ESMAMKTDVLTSPQSESMYKSLNHLITQSVASDSRNPFEHAIQYIAAHAT 108
D 73 DGEHWTF-----AAINAT 88
QY 109 EKDESILKHLIQGLDITILGSNDPFSYRNOIESAGLPLTPESLGLTAPFASITFNAES 168
D 89 -----TFKSEGLETFVKM-----VIEFLIDQ 112
QY 169 NGANSKEVAKTGLSSAAMTAVVAALL--HYLGVDLSDPCKEGFGCSLDLVIMIA 226
D 113 SG-----AKYGLSSAAMTAVVAALLMTKFEPIKMLK-----KLA 150
QY 227 QTSICLAQKVGSGFDVSCAVGSGQRYRFSPEVLSFAQVAVTGLPLNEVIGTILKKMD 286
D 151 ALSHLVYVNGSGCDIASCMYGMIAVTFPDQEWKHL-RLAKYSLE-----WKNKEWP 203
QY 287 NKRTFSLPLMLFLGEPGSGSSTPSMGAVKKM-QMSDEPKAREMNONLSDANLEET 346
D 204 MJOETLEEPVTPFSVGTGT-PVSTGKLVSOIHAFKQED-----SKNYQHL----- 251
QY 347 KLNDLSKLAQKHMVYLRVIRKSCSVLTSEKVLHATEPINEAIIKELLEAREAMLRIIL 406
D 252 RNEIMK-----QIDAFH--TKDELTLSSIKENRIIQL----- 286
QY 407 MRONGEASVPIPEESOTQLDSTMSAGVLAGVPGAGFCAIFATLDSGKLTQ 464
D 287 ---GTAGVAVIELSLKELAD---SAENMGAKSSSGGDCGAGFASFKELAEKLVN 339
QY 465 AW 466
D 340 EW 341

RESULT 6
AE1806
mevalonate kinases homolog lmo0012 [Imported] - Listeria monocytogenes (strain EGD-e)
C:Species: Listeria monocytogenes
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001
C:Accession: AE1806
R:Glaser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloec-
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurgeit, O.; Entian, K.D.; Fsihl,
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Krefit, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maltournam, A.; Ma-
ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehl-
A.; Title: Comparative genomics of Listeria species
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AE1806
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-359 <GLA>
A:Cross-references: GB:NC_003210; PIDN:CA098227.1; PID:g164409371; GSPDB:GN00177
A:Experimental source: strain EGD-e
C:Genetics:
A:Gene: lmo0012

Query Match	5.8%;	Score 149.5;	DB 2;	Length 359;
Best Local Similarity	20.5%;	Pred. No. 0.001;		
Matches	99;	Conservative	60;	Mismatches 146;
				Indels 177;
				Gaps 20;

```

QY 8 PGKVLMTGCVLYLVECPNAGNLVLTSTNARYAIVK-----PIIEEYK 48
Db 13 PGKLYVAGEVAVNESGHIATILTAVN-RYITLILEDSENNELMIPHYENPMPVGGELKP 71
QY 49 ESMAMKMTDVKLTSPQLSRESMTKLSLNHLTLOSVASDSRNPVEHAIOYIAAAHLAT 108
Db 72 DGEHWTF-----AAEINAT 87
QY 109 EKDKESLHKLLLOGDLITILGSNDPFYSYRNOIESAGLPETPESIGTLAPFASITPNAES 168
Db 88 -----TLKSGIELTPYKM-----IITELIDQ 111
QY 169 NGANSKPEVAKTGLSSSAAMTTAVVALL--HYLGVDLSDPCKEKGKGCSDLDVIMHA 226
Db 112 SG-----AKYGLSSAAATVAVAINALMTKFPYEISMUK-----KFLA 149
QY 227 QTSCHLOGKXKSGSEPDVSCAVYGSORVYRFSPEVLFSFQAVNTGLPMEVIGTILKCKWD 286
Db 150 ALSHLYVOVNGSCGDIASCMTGGMTAYTTPDOEWKH-RLAKYSLE-----WEKKEBP 202
QY 287 NKRTFESLPPLMNLFLGEPGSGSSTPSMGAVYKKWQMSDEKAEENQNTSDANLEJET 346
Db 203 MLOJETLEEPVPTFSVGWGT-PTVSTGKLVSQIHAFKED---SKNYOHFL-----T 250
QY 347 KLNDLSKIAKDHMDVYLRLVIKSCSVLTSEKNVLAHTEPINEALIKELLEAREAMLRITIL 406
Db 251 RNNEIMK-----OIQAFH--TKDEELYSAIKENRRITQEL----- 285
QY 407 MROMGEAASVPIEESQQLDLDSTMSAGVLLAG-VPGAGFDAIFAIT-LDGSCTKLTO 464
Db 286 ---GTRKGVNIETSLTELAD---SAENMGAGKSSSGGDDCICIANSKTELAKELVN 338
QY 465 AM 466
Db 339 EM 340

```

RESULT 7
B89828
phosphomevalonate kinase [imported] - *Staphylococcus aureus* (strain N315)
C:Species: *Staphylococcus aureus*
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C:Accession: B89828
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Ogura, M.; A.; Mitsuhashi, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.; C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of methicillin-resistant *Staphylococcus aureus*.
A:Reference number: A89758; MUID:21311952; PMID:11418146
A:Accession: B89828
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-358 <RUR>
A:Cross-references: CB:BA000018; PTD:913700484; PIDN:BBB41781.1; GSPDB:GN00149
A:Experimental source: strain N315
C:Genetics:
C:Gene: mwak2

Query Match	5 2#:	Score 134.5:	DB 2:	Length 358:
Best Local Similarity	21.8%:	Pred. No. 0.015:		
Matches 104:	Conservative 59:	Mismatches 181:	Indels 133:	Gaps 23:
OY	3 VVASAPGKVLMTGGYLVEKPAGLIVLTSTNAFFAIVKPIINEEVKPESSAMRWTDVILTS	62		
	: : :			:
Dd	2 IQVKAPGRLTYAGEVAITE-PQKSVLLALDLRF-----VT	35		
OY	63 POLSFESWKLISLNHLTQSASADSRNPFVHAIIQYAIATAAHNLTEKDESLIKILLILOG	122		
	: : :: : : : :			:

```

Db      36  ATEEAQKYKTHSKALHNHPVFSRD---EDSI--VIDPFAAKQ-----77
QY      123  LDITILGNDIFYSTRNOIES--GLPTPESLGLAPFATITPNAESGANSKPREVAT 180
Db      78  LNYVVALEIEFOYAKSCDIAKMHFLITDS-----NDDSGN-----HXY 118
QY      181  GLGSSAMTFAVVALAHYGVVDLSDPCKEKFGSGDDIDYHMAIOTSCHLACQKVGSG 240
Db      119  GLGSSAAVLVSIVLNEFDIM-----KSNL-YIKKLAVIANMKLOJLSGSG 165
QY      241  FDVSCAVY-GSQRVRESPELSEFAOYAVTGLPLNEVITIGLTKGWDNRKTE-FSLPPLM 298
Db      166  -DIASVYSGWLAVSTFDHEWKH-QLEDF-----TVEVLIKNPGHIEPLAPENM 217
QY      299  NLFTGEPGSGSGSTPSPWAGVAKKQMSDPKAEKNMNLSDANLEETKLNDLSKLAKH 358
Db      218  EYLIGMTGSPASS-PHEVSEVKRLK-SDP-----SF 246
QY      359  MDVLYLRIKCSVLTSEKWVLHTEPIENALIKELLEAREAMLRIILRMQGEASVPI 418
Db      247  YGDFLEDSHRC---VER-LIHAFTNNIKGVQAWRON-----RTIIRMDKEATVVI 295
QY      419  EPEQOTOLDLSTMSAEVLLAGVAGGAFPAIFATLGD-SGTKLQOAWSSHVAVL 474
Db      296  ETEKTKYICDLAEKYNHGA--SKTSGAGGCGGCTIINKQDKRKIYDEWTKHGIRPL 350

```

T21713
hypothetical protein F33H2.2 - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jan-2000
C:Accession: T21713
R:Collage, A.
Submitted to the EMBL Data Library, November 1996
A:Reference number: Z19463
A:Accession: T21713
A:Status: Preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-816 <WIL>
A:Cross-references: EMBL:Z81526; PIDN:CAB04264.1; GSPDB:GN00019; CESP:F33H2.2
A:Experimental source: Clone F33H2
C:Genetics:
A:Gene: CESP:F33H2.2
A:Map position: 1
A:Introns: 51/1, 101/3, 186/1, 270/3, 328/2, 380/3, 509/2, 751/2
C:Superfamily: *Caenorhabditis elegans* hypothetical protein F33H2.2

[illegible]

Db 356 ALSP-----VSEDSALRASSPNPSLNNSSVESFSLSTQFASDLSSGSGYNRAAFIFDS 410
QY 340 -----ANLEJETK-----LNDLSKLAKDMDVLRVIRKSCSVLT---SEKVVLATE 383
Db 411 TLTAFLMGLNLSLKGHAVALFEVCKLADEQMKLEQLESYVQFAEGDAQRMHATA 470
QY 384 PINEAIKELLEARL-----RIRILMRQK-----EASVPTEP 420
Db 471 LBS--LKSIRQERADLINGSISITLDEKSRQVLOKSGILVMAAPLSTECATIV-- 526
QY 421 ESQTQLDSTMSAEVGLVAGVPGAGFDALFAITLGDG-----GTLTLT----- 463
Db 527 -----QSVFLIPPNAPETCSPMFRLSIYSACSSGSPSVLPGLRILTLTPRV 573
QY 464 -QAMSSHNVALLVREDPH 481
Db 574 LQSFQHRHFLVSSTKHEPH 592

RESULT 9

F70185
phosphomevalonate kinase homolog - Lyme disease spirochete
C:Species: Borrelia burgdorferi (Lyme disease spirochete)
C:Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 08-Oct-1999
C:Accession: F70185
R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; White
son, D.; Peterson, J.; Kierlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vugt,
Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.
Nature 390, 580-586, 1997
A:Authors: Smith, H.O.; Venter, J.C.
A>Title: Genomic sequence of a Lyme disease spirochete, Borrelia burgdorferi.
A:Reference number: A70100; MUID:98065943; PMID:9403685
A:Accession: F70185
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-317 <RLE>
A:Cross-references: GB:AE001169; GB:AE000783; NID:g2688611; PIDN:AAC67030.1; PID:g2688611
A:Experimental source: strain B31

Query Match 4.88; Score 124.5; DB 2; Length 317;
Best Local Similarity 18.38; Pred. No. 0.072;
Matches 86; Conservative 65; Mismatches 141; Indels 177; Gaps 19;
QY 6 SAPGKVLMTGCVLLEKPNAGLVSTNARFYAIRKPIINEEVKPESMAMKMTDVLTPQL 65
Db 7 SVGNLLMEYTLLEKGLGALAIKRAFFSK-----KDSMRP-----F 49
QY 66 SRESMYKLSLNHLTLQSVASDSRNPFVEHAIOYAIAMHLATEKDESLHKLGLGIDI 125
Db 50 SKKK-----KIDFSLIENRSDVFKMFAY----- 74
QY 126 TLIGSNDFYRYNOIESAGLPPIESLGTAPASTIFPNAESNGANSKEPAVATGIGSS 185
Db 75 --LSQNFCEFLIEN-----FAVDYIDTSNPFEN-----DGTKKGGSS 110
QY 186 AAMTTAVVALL--HYLGVDLSDPCKEKGKSGSDLDVIMIAQTSKCL-----AGKV 237
Db 111 AVAIGVIGLFLIHNTNVEKEIFK-----YCELAIRYSOGGI 151
QY 238 GSGEDVSCAVYGSQRYRFSPEVLSPAQVAVTGLPLNEVIGTILKGMKDKRTFESLPL 297
Db 152 GSGVDIATSIFFG-----VIEF-----EGFNFKCRQLGAVEF 184
QY 298 MNLFLGPGSGSGSTPMGAVKKWMSDPEKARENMONSDANLELETUKLNDLSKLAKD 357
Db 185 NDFL-----MQLQ--AIKTTTICEYENK----- 207
QY 358 HMDVLYIKSVLTSEKVLHATEPINEAIKELLEARMRIRILMRQEAASVP 417
Db 208 HRNITLPIIKCN-LBKMKLVLNNSNS-KSALLSLRRARELGLAIT--EALGVSAALP 262
QY 418 IEPE--SQQLDSTMSAEVGLVAGVPGAGFD--AIFAITLGDGSGT 461

Db 263 SSFDHLIGQDLIKALGAGNETFLVYRPNIEAFNLKISIVLENIGIK 311

RESULT 10

S42088
mevalonate kinase (EC 2.7.1.36) - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: S42088
R:Riou, C.; Jourte, Y.; Lacroste, F.; Karst, F.
submitted to the EMBL Data Library, February 1994
A:Description: Isolation and characterization of a cDNA encoding Arabidopsis thaliana
A:Reference number: S42088
A:Accession: S42088
A:Molecule type: mRNA
A:Residues: 1-378 <RLO>
A:Cross-references: EMBL:X7793; NID:g456613; PIDN:CAA54820.1; PID:g456614
A:Function:
C:Description: catalyzes the reversible phosphorylation of (R)-mevalonic acid by ATP
A:Pathway: isoprenoid biosynthesis
A:Note: this enzyme can also utilize GTP, CTP, and UTP
C:Superfamily: mevalonate kinase
C:Keywords: ATP; isoprenoid biosynthesis; phosphotransferase
F:352/339/Region: ATP binding #status predicted
F:352/Binding site: ATP (Lys) #status predicted

Query Match 4.88; Score 123; DB 1; Length 378;
Best Local Similarity 20.98; Pred. No. 0.13;
Matches 100; Conservative 78; Mismatches 151; Indels 150; Gaps 23;

QY 1 MAVVASAPGKVLMTGCVLLEKPNAGLVSTNARFYAIRKPIINEEVKPESMAMKMTDVL 60
Db 1 MEVARAPGKILLAGEHAVVHG-----AVAAID-----LYTVTL 38
QY 61 TSPQLSRESMYKLSLNHLTLQSVASDSRN-PEVEHAIOYAIAMHLATEK--DKESLHK 117
Db 39 RFLPISAENNDRILQ--LKDLSLEFSWSIARIKEAIPDSSSTLCRSTPASCSEETLK- 94
QY 118 LLLQGLDITLIGSNDFYRYNOIESAGLPPIESLGT-LAPF-----ASTFN 164
Db 95 -----SIAYL-----VEEQNLPEKMWLSSGISTFMLYRIIGFNPAIVIN 137
QY 165 AAEANGANSKPEVAKTGIGSSAAMTTAVVALLHYLGVDLSDPCKEKGKGCSD--LDV 221
Db 138 SELPYG-----SELGSSAALCVALLTAAL--ASSISEKTRNGSSLDDETLDEL 184
QY 222 IHMIAQTSKCLAQKVGSGFDVSCAVYGSQRYRFSPEVLSPAQVAVTGLPLNEVIGTIL 281
Db 185 LNKWAFEGEKIIRGK-PSGIDNTVSAVGN-----MIKFCGETIRLQSNMPLRMLI 234
QY 282 KGRKDKRTFESLPLMNLFLGPGSGSGSTPMGAVKKWMSDPEKARENMONSDAN 341
Db 235 --TNTRV-----GRNTKALVSGVSOAIVNHPAMSVFAVDSIS 272
QY 342 LELETUKLNDLSKLAKDMDVLRVIRKSCSVLTSEKVLHATEPINEAIKELLEARPAML 401
Db 273 KELAAITQ-----SKDETSV-----ERIKELMEAMQGL 305
QY 402 RIRILMRQEAASVPPIEPSQTQLDSTMSAEVGLVAGVPGAGFDALFAITLGDGSGT 460
Db 306 -----LSMGVSHSIEAVILTYVHK--LVSKILTAGGGGCV--LTLPLPGT 348

RESULT 11

A12447
hypothetical protein al15137 [imported] - Nostoc sp. (strain PCC 7120)
C:Species: Nostoc sp.
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 30-Jun-2002
C:Accession: A12447
R:Kanevo, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Irigu
Nakazaki, N.; Shimo, S.; Sugimoto, M.; Takazawa, M.; Yasuda, M.; Tabata

DNA Res. 8, 205-213, 2001
 A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium *Anaerostipes*
 A:Reference number: AB1807; MUID:21595285; PMID:11759840
 A:Accession: A12447
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-721 <KUR>
 A:Cross-references: GB:BA000019; PIDN:BAW6836.1; PID:g17134275; GSPDS:GN00179
 A:Experimental source: strain PCC 7120
 C:Genetics:
 A:Gene: all5137

Query Match 4.6%; Score 117.5; DB 2; Length 721;
 Best Local Similarity 22.3%; Pred. No. 0.95;
 Matches 96; Conservative 59; Mismatches 136; Indels 139; Gaps 19;

29 LSNARPY-ATVPP-----INEEVKESAMKMTVVK----- 59
 25 LAMNNHYQALAPTRPKALLIGINDYKSSLSGLTVELOKELLYRFGQATDIL 84
 60 -LTPOLSRSMYKLSLNLTLQS-----VSASDSRNPVEHAIOYAAHMLATE 109
 85 TLTEQASREFIEAFLDLTKQAKPDYVVFHPSGYGTQLPVESGTLQNAL---VTTD 140
 110 KDKESLHK-----LLLOGL-DITILGNDFYSYRNOIESAGLPL--TPESLG 153
 141 ENQADDSQIANYLLEDTLTLRLSLPTDHALVLDTSYFTGTINOPAGIKIRARESPG 200
 154 TLAPFASIFPN-----AASNGANSKPEVAKTGLG----- 183
 201 TRIAAEIDPROQIKNONPEFSPVILSATSDQOSAREILMSGFSGGLTYALTQOME 260
 184 SSAAMTTAV---VAALLHYLVVDLSDP---CKEGKFGSGDDLVIMHMTOSHCLAQK 236
 261 STPATIRVSLSHASSIHQLG--SKQPGILTKKNGQVATVENLLDRT----- 310
 237 VSGGFVSCVAGSQRVRSPEVLSPAQVAVGLPLNEVIGTILKGMKMKTEESLP 296
 311 -----GAGGATIAIEDGKTAQVMDGLP---AOVLENYGNSR--FTLAT 351
 297 LAMNLFGEPSGSGSTPSMVGAVKQMSDPERA-----RENNQNLSDANLEL 344
 352 GEELVLRSLMAGRS--PSRGLAKAQFSKVTETIPVQGLQIENVRVLSNINILAL 409
 345 ETKLNDLSKL 354
 410 DTKLERIERV 419

RESULT 12

mevalonate kinase (EC 2.7.1.36) - human
 N:Alternate names: MK
 C:Species: Homo sapiens (man)
 C>Date: 04-Mar-1993 #sequence_revision 18-Mar-1997 #text_change 11-Jun-1999
 C:Accession: A42919; A58527
 R:Schaefer, B.L.; Bishop, R.W.; Kratunis, V.J.; Kalinowski, S.S.; Mosley, S.T.; Gibson, R.J. Biol. Chem. 267, 13229-13238, 1992
 A:Title: Molecular cloning of human mevalonate kinase and identification of a missense mutation
 A:Reference number: A42919; MUID:92317034; PMID:1377680
 A:Accession: A42919
 A:Molecule type: mRNA
 A:Residues: 1-396 <SCH>
 A:Cross-references: GB:M88468; NID:g307197; PIDN:AAB59362.1; PID:g187561
 A:Note: sequence extracted from NCBI backbone (NCBIIN:107743, NCBIIP:107744)
 R:Graef, E.; Caselmann, W.H.; Wells, J.; Koshy, R.
 Oncogene 9, 81-87, 1994
 A:Title: Insertional activation of mevalonate kinase by hepatitis B virus DNA in a human
 A:Reference number: S42226; MUID:94134441; PMID:8302606
 A:Accession: A58527
 A:Molecule type: mRNA
 A:Residues: 1-396 <GRA>
 A:Cross-references: EMBL:X75311; NID:g450345; PIDN:CAA53060.1; PID:g450347

C:Genetics:

A:Gene: GDB:MWK
 A:Cross-references: GDB:134189; OMIM:251170
 A:Map position: 12pter-12qter
 A:Note: defects in this gene can cause mevalonic aciduria
 C:Function:
 A:Description: catalyzes the reversible phosphorylation of (R)-mevalonic acid by ATP
 A:Pathway: Isoprenoid biosynthesis
 A:Note: this enzyme can also utilize GTP, CTP, and UTP
 C:Superfamily: mevalonate kinase
 C:Keywords: ATP; isoprenoid biosynthesis; phosphotransferase
 F:335-357/Region: ATP binding #status predicted
 F:357/Binding site: ATP (lys) #status predicted

Query Match 4.5%; Score 117; DB 1; Length 396;
 Best Local Similarity 20.2%; Pred. No. 0.39;
 Matches 104; Conservative 74; Mismatches 183; Indels 154; Gaps 22;

3 VASAPGKVLMTGTYLVLEKPNAGVLSINARFYALVPLNEVKEPSNMMKMTDYKLS 62
 6 LTVSAPGKVLHGEHVVH-GKVALAVSLNLRFLRLQPSHN-----GKVDLSL 53
 63 POLSRSMYKLSLNLTLQSVASDSRNPVEHAIOYAAHMLATEKDKESLHKLLQG 122
 54 PNIGIKRAMDVA---RLQSLDTS-----FLEGD-----VTPPSQVEYKELK--- 92
 123 LDITILGNDFYSYRNOIESAGLPL--LTPESIGTLAPASITFMAESNGANSKEVA- 178
 93 -----EVALGLPDDCAVTERLAVLA-FLYVLISICRQORALPSLDIV 133
 179 -----KTGLGSSAMTTAVVAALLHYLVVDLSDPCKG---KKGCDLVYIMHIAQT 228
 134 WSELPPGAGIGSSAAVSVCLAAALLTV--CEELPNLKGDCVNRWTKEDLEINMAQO 191
 229 SHCLAQKVGSGGFVSCVAGSQRVRSPEVLSPAQVAVGLPLNEVIGTILKGMKMK 288
 192 GEMHIGN-PSGVDNVASTMG-----ALRYQKGISLSKRPAIOILL----- 234
 289 RFEESLPLMLNLFGEPSGSGSTPSMVGAVKQMSDPERA-----RENNQNLSDANLEL 348
 235 -TWTKP-----RNTRALVAGVRNRLKFEIYAPLTSIDALSICEERYL 279
 349 NDISKLAKHMDVYLVAVINSCSVLTSEKWLHATEPINEATIKELLEARAMRILMR 408
 280 GEMGE-----APAEQYLVLEELIDMMQHILN----- 306
 409 QMGEASVPIEPESQTLDDSTMSAGVLLAGVPGAGFDALFATLGDSTKLTQAMSS 468
 307 -----ALGVGHASLDQLQVLT-RARG-LHSKLTGAGGGGC-GITLLKPGLEQPEVAT 356
 469 HNVALLVREDPHGVCLESQDPPTTCITSGVSSIH 503
 357 KQALTCGFD-----CLE-----TSIGAPGVSIH 380

RESULT 13

S42226
 hypothetical large surface antigen/mevalonate kinase (EC 2.7.1.36) mutant fusion prot
 C:Species: hepatitis B virus, HBV
 C>Date: 13-Jan-1995 #sequence_revision 17-Mar-1997 #text_change 20-Oct-2000
 C:Accession: S42226
 R:Graef, E.; Caselmann, W.H.; Wells, J.; Koshy, R.
 Oncogene 9, 81-87, 1994
 A:Title: Insertional activation of mevalonate kinase by hepatitis B virus DNA in a hu
 A:Reference number: S42226; MUID:94134441; PMID:8302606
 A:Accession: S42226
 A:Molecule type: mRNA
 A:Residues: 1-557 <GRA>
 A:Cross-references: EMBL:X75311; NID:g450345; PIDN:CAA53059.1; PID:g450346
 C:Keywords: phosphotransferase
 F:1-157/Region: hepatitis B virus large surface antigen (fragment)
 F:162-557/Region: human mevalonate kinase (fragment)

	Query Match	4.5%	Score 116;	DB 2;	Length 335;
	Best Local Similarity	21.9%;	Pred. No. 0.36;		
	Matches	102;	Conservative	60;	Mismatches 124; Indels 180; Gaps 25.
OY	3	VVASAPGVLMTGY-LVLEKPNAGLVLTSTNARFAIVKPINE-EVARESNAMRWTD--	57		
	: :	:	: :		
D6	5	VLASAPAVITLGEHSVVYGR-AIAAIEIERTY-VRAQFDSNCIKIEADIIKTPLGI	61		
	: :	:	: :		
OY	58	VKLSPQLSRSEMYLSTLNHLTGVSASDSDNPVEVAIQVALAAHLLATFKDKESLHK	117		
	: : : :	:	: :		

Dd	62	VSFSDKTIYFEDDYCKAAEVLST-----VYRAIELAL-----BEDSKR-----	99
Oy	118	LILGLDITLIGSNDFYSRNQIESAGLPPTPESTGLTAPASTIFENNAESGANSKEPV	177
		:::	
Dd	100	--VGIDVISI-----TSQIPV-----GA-----	114
Oy	178	AKTGIGSSAAMTTAVVAALLHYLGAVDLSDPCKEKFCGSDLDYHMTAQSH---CLAQ	234
		:::::	::: :
Dd	115	--GGSSAAAVATATGAVSRLLG-LELS--KE-----ETAKIGHVEYLQ	155
Oy	235	GKVSGGFVSCAVYSQRXY---RSPSEVLSFAQAVALGPLNEVIYGLTKGMKNKRTE	291
		:::	:
Dd	156	G-ASSGIDPYTSAVNGCFLYKKQKREP-----LPPME-----	186
Oy	292	FSLPPLMNLFLGEPSGGSSSTPSMYGAVKMKOMSDPEKARENMONTLANELEFKINDL	351
		:::	: :
Dd	187	--LP-----IVGYGTSTG-STKELVAMVRKRYEEPM-----LVPILIFAMGKLVDK	231
Oy	352	SKLAADHDMDYLIRVLIKSSCVLTSEKXVLIHAATEPINELAIKELLEKREMLRIRIMR-QM	410
		:::	:::
Dd	232	AK-----EDILSKLEEEXLTRLGELMINIH	257
Oy	411	GEAAPVIEPIESPOTLIDSTMSAEGLLAGVYGAGCPAIFAFTLG	456
		:::	:::
Dd	258	GLLDALGVSTKKLGELYTAART-AIGAIAKTLAGGGCGMTALAPG	302
 RESULT 15 B86705			
ATP-dependent proteinase ATP-binding subunit [imported] - Lactococcus lactis subsp. 1			
C:Species: Lactococcus lactis subsp. lactis			
C:Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001			
C:Accession: B86705			
R:Boilotin, A.; Winker, P.; Manger, S.; Jallou, O.; Malarme, K.; Weissenbach, J.; Ehr			
genome Res. 11, 731-753, 2001			
A>Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis			
A:Reference number: AB6625; MUID:21235186; PMID:11337471			
A:Accession: B86705			
A>Status: Preliminary			
A:Molecule type: DNA			
A:Residues: 1-816 <STOD>			
A:Cross-references: GB:AB005176; PID:g12723546; PIDN:AAK04740.1; GSPDB:GN00146			
A:Experimental source: Strain IL1403			
C:Genetics:			
A:Gene: ClpC			
C:Superfamily: endopeptidase Clp ATP-binding chain			
Query Match 4.5% Score 115; DB 2; Length 816; Best Local Similarity 21.4%; Pred. No. 1.8; Matches 94; Conservative 68; Mismatches 164; Indels 114; Gaps 19;			
Oy	92	FVEHHIOYAIAAHAAHTKDKRESL---HKL-LLO-----GLDITIIGSNDFYSRNQIE	141
		:::	:::
Dd	82	FSPRAEVMYTAASPLAIIHNNSEAVSTEHLLYALLDVEDGPFQLTKLOKINIVSLRKELE	141
Oy	142	S-AGLPPTPESTGLTAPPA-SITFNAAESNG-----ANSKPENVATGIGSSAAMTTA	191
		:::	:::
Dd	142	KRTGKY-PESKKAATPMSKRMAKGVAENSTPTLDSVSDDLTEEARLGLDPMIGREA	200
Oy	192	VVAALLHYLGAVDLSDPCKEKFCGSDLDYHMTAQTHCLCAOKRGVSGFPVSCAVYGSQ	251
		:::	:::
Dd	201	EIDRIHLISRRTKNPVLVBEPGKSAITLEGAAOR---IVNSOVPRLG-----	247
Oy	252	RVRSPEVLSFAQAVALGPLNEVIYGLTKGMKNKRT-----EFSLPPLMNLFGE---	304
		:::	:::
Dd	248	---ANSMALNNMATV-----VAGTRFGFEFDRLTAIVEEYSSPDVYITRIDELHT	296
Oy	305	P-GSGGSSTP-----SMGA-----VKRWMSDPEKARENONT	337
		:::	:::
Dd	297	IIGAGGDMDSVDAANIILKPALARGDFQMVGATYYHEYOYIERDEALERLATINVIDEP	356
Oy	338	SD---AALE-LERKINDLSKLAADHDMDYLIRVLIKSSCVLTSEKXVLIHAATEPINEA----	388
		:::	:::

Db 357 SPDEAIAILOGLEKFEEDYHQVFTDOAIKSAVMSVRYMTSRKLPDKAIDLDEAAAAY 416
QY 389 -----IKELLEAR-----EAMLRIRIIMROMGEAA-----SVPIEPES 422
Db 417 KISVKNQOTKRLDLEKELTFEAOEELSEAVIKLDIKASRTKEKAVERKADKIYKFSYKEDK 476
QY 423 QTOLLSTMSAEGVLLAGVP 442
Db 477 ROEVTDOAVVAVASTLTGVP 496

Search completed: April 26, 2003, 12:57:30
Job time : 26 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 26, 2003, 12:50:50 ; Search time 14 Seconds

(without alignments)
1496.110 Million cell updates/sec

Title: US-09-988-863a-2
Perfect score: 2581
Sequence: 1 MAVVASAPKGVLTMTGTYLV.....ESGDPPTCTISGVSSIHLE 505

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues
Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : SwissProt.40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	453.5	17.6	451	1	ERG8_YEAST
2	123	4.8	378	1	KIME_ARATH
3	117	4.5	396	1	KIME_HUMAN
4	116	4.5	335	1	KIME_PYRHO
5	113.5	4.4	395	1	KIME_MOUSE
6	112	4.3	287	1	DHIL_RAT
7	112	4.3	335	1	KIME_PIRAB
8	111.5	4.3	395	1	KIME_RAT
9	109.5	4.2	292	1	DHIL_MOUSE
10	109	4.2	662	1	FTSH_CAPAN
11	107	4.1	529	1	NXIA_MOUSE
12	107	4.1	1442	1	VJ9F_YEAST
13	106	4.1	698	1	ALT_YEAST
14	103.5	4.0	284	1	KIME_ARCFU
15	103	4.0	332	1	LDHA_BRABE
16	103	4.0	504	1	GALL_CANPA
17	103	4.0	710	1	GSOD_ERMCH
18	101	3.9	475	1	PBR_SERMA
19	100	3.9	292	1	DHIL_HUMAN
20	100	3.9	698	1	ALT_BPT2
21	100	3.9	1158	1	RIL4_HUMAN
22	99.5	3.9	443	1	KIME_YEAST
23	99.5	3.9	463	1	YDJN_ECOLI
24	99.5	3.9	1066	1	PKI1_YEAST
25	98	3.8	2493	1	CYAA_USTMA
26	97	3.8	292	1	DHIL_SHEEP
27	96.5	3.7	505	1	FIXL_BRAJA
28	96.5	3.7	714	1	FTSH_TOBAC
29	96	3.7	854	1	DISL_HUMAN
30	96	3.7	1163	1	SBCC_CIOAB
31	95	3.7	331	1	LDHA_MACRA
32	95	3.7	893	1	GEFH_HUMAN
33	95	3.7	3432	1	POLG_JAEV1

34	94.5	3.7	400	1	YXXA_BACSU	P39668 bacillus su
35	94.5	3.7	735	1	IF2C_GUTH	C78489 guillardia
36	94.5	3.7	2841	1	NFL_MOUSE	O04690 mus musculu
37	94	3.6	618	1	MBHL_RHOE	P17632 rhodocyclu
38	94	3.6	1209	1	DNBI_HSVB	P28932 equine herp
39	94	3.6	3432	1	POLG_JAEV5	P19110 j genome po
40	93.5	3.6	515	1	GALL_CANLA	P56091 candida alb
41	93.5	3.6	1157	1	XYNA_THESA	P36917 thermoaer
42	93	3.6	712	1	GSOD_ERMCH	P31700 erwina chr
43	93	3.6	730	1	ER2_METH	027131 methanobact
44	93	3.6	763	1	ECHA_PIG	Q29534 sus scrofa
45	92.5	3.6	922	1	DPOL_RICPE	O9raa9 rickettsia

ALIGNMENTS

```

RESULT 1
ERG8_YEAST          STANDARD:      PRT: 451 AA.
ID                  ID
AC P24521:
DI 01-MAR-1992 (Rel. 21, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Phosphomevalonate kinase (EC 2.7.4.2).
GN ERG8 OR YMR220W OR YM9959.02.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE OF 1-424 FROM N.A.
RX MEDLINE=91117228; PubMed=1846667;
RA Tsay Y.H., Robinson G.W.;
RT Cloning and characterization of ERG8, an essential gene of
RT Saccharomyces cerevisiae that encodes phosphomevalonate kinase.*;
RL Mol. Cell. Biol. 11:620-631(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S288c / AB972;
RA Skelton J., Churcher C.M., Barrell B.G., Rajandream M.A., Walsh S.V.;
RA Submitted (JUN-1995) to the EMBL/Genbank/DBJ databases.
CC -! FUNCTION: ERG8 IS AN ESSENTIAL GENE.
CC -! CATALYTIC ACTIVITY: ATP + (R)-5-phosphomevalonate = ADP + (R)-5-
CC -! PATHWAY: SECOND STEP IN ISOPENTENYL DIPHOSPHATE FORMATION.
CC -! SUBCELLULAR LOCATION: Cytoplasmic.
CC -! SIMILARITY: BELONGS TO THE GMP KINASE FAMILY. MEVALONATE KINASE
CC -! SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
CC EMBL, M63648; AAA34596.1; -
CC EMBL, Z49939; CA90191.1; -
CC PIR, A39606; A39606.
CC SGD, S0004833; ERG8.
CC InterPro: IPR001745; GHMPKase_ATP.
CC TIGRfam: TIGR01219; Pmev_kin_ERG8; 1.
CC PROSITE, PS00627; GHMP_KINASES_ATP; 1.
CC Transferrase; kinase; ATP-binding; isoprene biosynthesis;
CC sterol biosynthesis.
CC NP_BIND 150 160 ATP (POTENTIAL).
CC BINDING 183 183 ATP (POTENTIAL).
CC CONFLICT 213 213 A->R (IN REF. 1).
CC CONFLICT 418 423 TANDR -> PLMTND (IN REF. 1).
CC SEQUENCE 451 AA; 5045 MM; DB2B6C862151683C CRC64;

```

Query Match 17.6%; Score 453.5; DB 1; Length 451;
 Best Local Similarity 30.1%; Pred. No. 1.8e-26;
 Matches 155; Conservative 72; Mismatches 185; Indels 103; Gaps 18;

6 SAAGKVLMTGCVLEKPNAGLVSTNARFYAIYKPIINE-----EVRKESNAWK---W 55
 8 SAAGKALLAGGYLVLDKYEAFVWGSARMVAHAHPYSGLSGSDKEFEVRKSKQFKGEW 67
 56 TDVKLTSPOLSRSMYKLSLNHTLOS-----VSASDSRNPFEVHAIOYATAAHLATEKD 111
 68 -----LYHISPKSGFLPVISIGSGSKNFIKLVIANVSYF----- 101
 112 KESLHKLLOGDITLIGSDFTSYRNOISAGLPLTPESLGLAPASITFNAASNGA 171
 102 KPMMDYCNMNLFEVDFSD--AYHSQEDS-----VTEHRG-----NRLSFHS 144
 172 NSRPEVAKTGLGSSAAMTAVVAALLHYLVLDSPCKGKGCSDLDVYHMAQTSHC 231
 145 HRIEVPKTLGSSAGLVLTALASFF-VSDLENNVDYR-----EYIHNAQVAHC 197
 232 LAQKVGSGFVSCAVYSGQRYVRFSPVLSPAQAVTGLPLNEVIGTILKGR----- 284
 198 QAQKIGSGFDVAAGSIRYRFPALLS-----NLP--DISATYGSKLHLVD 247
 285 ---WDKRTFSLPLMLNLFLEPGSGSSTPSKVGAVKKWQSDPEKARNMONLSDAN 341
 248 EEDMNTIKSNHLPSCGLTMGDI-KNGSETVKLVQKVMYDSHMPESLKITELDHAN 306
 342 LELETKINDLSKLAKHDWYLVLRKSC--SVLTSEKWLHATPEINAIKELLEAREA 399
 307 SRMDGLSKLDRLHETHDDSDQIFESLERNDCTCKY-----PETTERDA 353
 400 MLRIILRMQGEAAVYPIEPESQTLDDSTMAEVLGAVPGAGFDAIFAITLGDSC 459
 354 VATIRSFRTITESGADIPVQTSLLDDCQTLKGLVLTCLIPAGGYDAIAVITKQVD 413
 460 TKLTQAMSSH-NVALLVREDPHGVCLESGDPT 492
 414 LRAQTANDKRFKRYOMLDVTOADWGVKKEK-DBET 447

DB 414 LRAQTANDKRFKRYOMLDVTOADWGVKKEK-DBET 447

RESULT 2
 KINE_ARATH STANDARD; PRT; 378 AA.

ID KINE_ARATH STANDARD; PRT; 378 AA.
 AC P46086;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Mevalonate kinase (EC 2.7.1.36) (MK).
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.
 NCBI_TaxID:3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20252529; PubMed=10794536;
 RA Lluich M.A., Masferrer A., Arro M., Boronat A., Ferrer A.;
 RT "Molecular cloning and expression analysis of the mevalonate kinase
 gene from Arabidopsis thaliana.";
 RL Plant Mol. Biol. 42:365-376(2000).
 CC -1- CATALYTIC ACTIVITY: ATP + (R)-mevalonate -> ADP + (R)-5-
 phosphate mevalonate.
 CC -1- ENZYME REGULATION: ITS ACTIVITY IS INHIBITED IN VITRO BY GERANYL
 PYROPHOSPHATE (GPP) AND FARNESYL PYROPHOSPHATE (FPP) THAT BIND

COMPETITIVELY AT THE ATP-BINDING SITE ON THE ENZYME.
 -1- PATHWAY: CHOLESTEROL BIOSYNTHESIS, MEVALONATE CATABOLISM.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Potentially).
 CC -1- SIMILARITY: BELONGS TO THE GHMP KINASE FAMILY. MEVALONATE KINASE
 SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL: X77793; CAA54820.1; -
 DR EMBL: AF141853; AAD31719.1; -
 DR InterPro: IPR001745; GHMPkinase_ATP.
 DR InterPro: IPR001459; Mev_gal_kin.
 DR Pfam: PF00288; GHMP_kinases.1.
 DR PRINTS: PR00959; MEVGALKINASE.
 DR TIGRFAMs: TIGR00549; mevalon_kin; 1.
 DR PROSITE: PS00627; GHMP_KINASES_ATP; 1.
 KW Transferase; Kinase; Sterol biosynthesis; ATP-binding.
 FT NP_BIND 141 151
 SQ SEQUENCE 378 AA; 40643 MW; 87E7E5C78BE53C66 CRC64;

Query Match 4.8%; Score 123; DB 1; Length 378;
 Best Local Similarity 20.9%; Pred. No. 0.078;
 Matches 100; Conservative 78; Mismatches 151; Indels 150; Gaps 23;

1 MAVVASAPGVMTGCVLEKPNAGLVSTNARFYAIYKPIINEEVRKESNAWKMDVKL 60
 1 MEKARAPGKIIILGHAHVHG-----ST-----AVAAID-----LYTVTL 38
 61 TSPQLSRSMYKLSLNHTLOSASASDRN-PRVEHAIOYATAAHLATEK--DKESLHK 117
 39 RFPPLSAENDRLTLQ--LKDLSLEFSWGLARIKEAIPDSSILCRSTPASCSEETLK 94
 118 LLLQGLDITLIGSDFTSYRNOISAGLPLTPESLGL-LAPF-----ASITFN 164
 95 -----SIATL-----VEQNLPRKEMWLLSSGISITFLMLYRIIGFNPATVIV 137
 165 AAEISGANSKPEVAKTGLGSSAAMTAVVAALLHYLVLDSPCKGKGCSD--LDV 221
 138 SELPYG-----SGIGSSAALCVALTAL--ASSISEKTRNGMSSIDETVLEL 184
 222 IHMAQTSHLAQKVGSGDVSCAVYSGQRYVRFSPVLSPAQAVTGLPLNEVIGTIL 281
 185 LNKWAFPEGEKTIHKK-PSGIDNTVSAVGN-----MKFSGEITRLOSMPMLMLI 234
 282 KGWMDKRTFESLPLMLNLFLEPGSGSSTPSKVGAVKKWQSDPEKARNMONLSDAN 341
 235 -----TNRV-----GRMTKALYSVSQRAVRHPRAMKSVFNAVDIS 272
 342 LELETKINDLSKLAKHDWYLVLRKSCSVLTSEKWLHATPEINAIKELLEBARAML 401
 273 KEIAAIIIO--SKDETSAV-----TEK-----EERIKELMEMONGLT 305
 402 RIRILMQGEAAVYPIEPESQTLDDSTMAEVLGAVPGAGFDAIFAITLGDSDPT 460
 306 -----LSMGVSHSIEAVILTVKHK-LVSKLTGAGGGCV-LTLLPPTGT 348

RESULT 3
 KINE_HUMAN STANDARD; PRT; 396 AA.

ID KINE_HUMAN STANDARD; PRT; 396 AA.
 AC Q03426;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Mevalonate kinase (EC 2.7.1.36) (MK).
 GN MK.
 OS Homo sapiens (Human).

CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A., AND VARIANT THR-301.
 RX MEDLINE=92317034; PubMed=1377680;
 RA Schaefer B.L., Bishop R.W., Kratunski V.J., Kalinowski S.S.,
 RA Mosley S.T., Gibson K.M., Tanaka R.D.;
 RT "Molecular cloning of human mevalonate kinase and identification of a
 RT missense mutation in the genetic disease mevalonic aciduria.";
 RL J. Biol. Chem. 267:13229-13238(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Hepatosoma;
 RX MEDLINE=94134441; PubMed=8302606;
 RA Graef E., Caselmann W.H., Wells J., Koshy R.;
 RT "Insertional activation of mevalonate kinase by hepatitis B virus DNA
 RT in a human hepatoma cell line.";
 RL Oncogene 9:81-87(1994).
 RN [3]
 RP SEQUENCE FROM N.A., VARIANTS HIDS P-20; P-39; L-135; T-148; T-268 AND
 RP I-377, AND VARIANTS MEVALONICACIDURIA P-20; F-264; T-268; T-334 AND
 RP M-310.
 RX MEDLINE=21214737; PubMed=11333768;
 RA Houten S.M., Koster J., Romeijn G.-J., Frenkel J., Di Rocco M.,
 RA Caruso U., Landrieu P., Kelley R.I., Kuis W., Poll-The B.T.,
 RA Gibson K.M., Wanders R.J.A., Waterham H.R.;
 RT "Organization of the mevalonate kinase (MK) gene and identification
 RT of novel mutations causing mevalonic aciduria and
 RT hyperimmunoglobulinemia D and periodic fever syndrome.";
 RL Eur. J. Hum. Genet. 9:253-259(2001).
 RN [4]
 RP VARIANTS MEVALONICACIDURIA ILE-243; PHE-264; PRO-265 AND THR-268.
 RX MEDLINE=99347937; PubMed=10417275;
 RA Hinson D.D., Ross R.W., Kriess S., Shaw J.L., Kozich V.,
 RA Rolland M.-O., Divry P., Mancini J., Hoffmann G.F., Gibson K.M.;
 RT "Identification of a mutation cluster in mevalonate kinase deficiency,
 RT including a new mutation in a patient of Mennonite ancestry.";
 RL Am. J. Hum. Genet. 65:327-335(1999).
 RN [5]
 RP VARIANTS MEVALONICACIDURIA MET-310 AND THR-334.
 RX MEDLINE=99330561; PubMed=10401001;
 RA Houten S.M., Romeijn G.-J., Koster J., Gray R.G.F., Darbyshire P.,
 RA Smit G.P.A., de Klerk J.B.C., Duran R., Gibson K.M., Wanders R.J.A.,
 RA Waterham H.R.;
 RT "Identification and characterization of three novel missense mutations
 RT in mevalonate kinase cDNA causing mevalonic aciduria, a disorder of
 RT isoprene biosynthesis.";
 RL Hum. Mol. Genet. 8:1523-1528(1999).
 RN [6]
 RP VARIANTS HIDS PRO-20; THR-268 AND ILE-377.
 RX MEDLINE=99295935; PubMed=10369261;
 RA Houten S.M., Kuis W., Duran R., de Koning T.J., van Royen-Kerkhof A.,
 RA Romeijn G.-J., Frenkel J., Dorland L., de Barse M.M.J.,
 RA Huijbers W.A.R., Rijkers G.T., Waterham H.R., Wanders R.J.A.,
 RA Poll-The B.T.;
 RT "Mutations in MK, encoding mevalonate kinase, cause
 RT hyperimmunoglobulinemia D and periodic fever syndrome.";
 RL Nat. Genet. 22:175-177(1999).
 RN [7]
 RP VARIANTS HIDS LEU-167; THR-268 AND ILE-377.
 RX MEDLINE=99295936; PubMed=10369262;
 RA Drenth J.P.H., Cuisset L., Grateau G., Vasseur C.,
 RA van der Velde-Visser S.D., de Jong J.G.N., Beckmann J.S.,
 RA van der Meer J.W.M., Delpech M.;
 RT "Mutations in the gene encoding mevalonate kinase cause hyper-IgD and
 RT periodic fever syndrome.";
 RL Nat. Genet. 22:178-181(1999).
 RN [8]
 RP VARIANTS HIDS N-20; P-39; L-150; L-167; R-202; O-215; T-268;
 RP S-309; R-326 AND I-377, VARIANT MEVALONICACIDURIA T-334, AND VARIANT
 RP N-52.
 RX MEDLINE=21214738; PubMed=1133769;

RA Cuisset L., Drenth J.P.H., Simon A., Vincent M.F.,
 RA van der Velde-Visser S.D., van der Meer J.W.M., Grateau G.,
 RA Delpech M.;
 RT "Molecular analysis of MK mutations and enzymatic activity in
 RT hyper-IgD and periodic fever syndrome.";
 RL Eur. J. Hum. Genet. 9:260-266(2001).
 CC -I- FUNCTION: MAY BE A REGULATORY SITE IN CHOLESTEROL BIOSYNTHETIC
 CC PATHWAY.
 CC -I- CATALYTIC ACTIVITY: ATP + (R)-mevalonate = ADP + (R)-5-
 CC phosphomevalonate.
 CC -I- ENZYME REGULATION: FARNESYL- AND GERANYL-PYROPHOSPHATES ARE
 CC COMPETITIVE INHIBITORS.
 CC -I- PATHWAY: CHOLESTEROL BIOSYNTHESIS, MEVALONATE CATABOLISM.
 CC -I- SUBUNIT: HOMODIMER.
 CC -I- SUBCELLULAR LOCATION: Cytoplasmic and peroxisomal.
 CC -I- DISEASE: DEFECTS IN MK ARE THE CAUSE OF MEVALONICACIDURIA. IT IS
 CC AN ACCUMULATION OF MEVALONIC ACID WHICH CAUSE A VARIETY OF
 CC SYMPTOMS SUCH AS PSYCHOMOTOR RETARDATION, DYSMORPHIC FEATURES,
 CC CATARACTS, HEPATOSPLENOMEGALY, LYMPHADENOPATHY, ANEMIA, HYPOTONIA,
 CC MYOPATHY, AND ATAXIA.
 CC -I- DISEASE: DEFECTS IN MK ARE THE CAUSE OF HYPERIMMUNOGLOBULINEMIA D
 CC AND PERIODIC FEVER SYNDROME (HIDS); AN AUTOSOMAL RECESSIVE DISEASE
 CC CHARACTERIZED WITH SKIN RASH, DIARRHEA, ADENOPATHY (SWOLLEN, TENDER
 CC LYMPH NODES), AND ATRALGIAS AND/OR ARTHRITIS. CONCENTRATION OF
 CC IGD, AND OFTEN IGA, ARE ABOVE NORMAL.
 CC -I- SIMILARITY: BELONGS TO THE GMP KINASE FAMILY. MEVALONATE KINASE
 CC SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M88468; AAB59362.1; -;
 DR EMBL: X75311; CAAS3060.1; -;
 DR EMBL: X75311; CAAS3059.1; ALT_INT.
 DR EMBL: AF217535; AAF82407.1; -;
 DR EMBL: AF217528; AAF82407.1; JOINED.
 DR EMBL: AF217529; AAF82407.1; JOINED.
 DR EMBL: AF217530; AAF82407.1; JOINED.
 DR EMBL: AF217531; AAF82407.1; JOINED.
 DR EMBL: AF217532; AAF82407.1; JOINED.
 DR EMBL: AF217533; AAF82407.1; JOINED.
 DR EMBL: AF217534; AAF82407.1; JOINED.
 DR PIR: A42919; A42919.
 DR GeneW: HGNC:7530; MKV.
 DR MIM: 251170; -;
 DR MIM: 260920; -;
 DR InterPro: IPR001745; GMPKase_ATP.
 DR InterPro: IPR001459; Mv_gal_kin.
 DR Pfam: PF002088; GMP_kinases.1.
 DR PRINTS: PR00959; MEVALKINASE.
 DR TIGRFAMs: TIGR00549; mevalon_kin.1.
 DR PROSITE: PS00627; GMP_KINASES_ATP.1.
 KW Transferase; Kinase; Cholesterol biosynthesis; ATP-binding;
 KW Peroxisome; Disease mutation; Polymorphism.
 KW NP_BIND 138 148
 FT FT
 FT VARIANT 20 20
 FT H -> P (IN HIDS AND MEVALONICACIDURIA).
 FT /FTID=VAR_004022.
 FT VARIANT 39 39
 FT L -> P (IN HIDS).
 FT /FTID=VAR_010957.
 FT VARIANT 52 52
 FT S -> N.
 FT /FTID=VAR_010958.
 FT VARIANT 135 135
 FT S -> L (IN HIDS).
 FT /FTID=VAR_010959.
 FT VARIANT 148 148
 FT A -> T (IN HIDS).
 FT /FTID=VAR_010960.

```

FT  VARIANT  150  150  S -> L (IN HIDS).
FT  VARIANT  167  167  /FTID=VAR.010961.
FT  VARIANT  202  202  P -> L (IN HIDS).
FT  VARIANT  215  215  /FTID=VAR.004023.
FT  VARIANT  243  243  G -> R (IN HIDS).
FT  VARIANT  264  264  /FTID=VAR.010962.
FT  VARIANT  265  265  R -> Q (IN HIDS).
FT  VARIANT  266  266  /FTID=VAR.010963.
FT  VARIANT  301  301  T -> I (IN MEVALONICACIDURIA).
FT  VARIANT  309  309  L -> F (IN MEVALONICACIDURIA).
FT  VARIANT  310  310  L -> P (IN MEVALONICACIDURIA).
FT  VARIANT  326  326  I -> T (IN HIDS AND MEVALONICACIDURIA).
FT  VARIANT  334  334  /FTID=VAR.004025.
FT  VARIANT  377  377  N -> T (IN MEVALONICACIDURIA; DIMINISHED ACTIVITY).
FT  VARIANT  396  396  /FTID=VAR.004026.
FT  VARIANT  424  424  V -> I (IN HIDS; MOST FREQUENT MUTATION).
FT  VARIANT  441  441  /FTID=VAR.004027.
FT  VARIANT  458  458  SQUOENCE 396 AA; 42451 MW; C8F6B629B58CD229 CRC64;

```

Query Match 4.5%; Score 117; DB 1; Length 396;
 Best Local Similarity 20.2%; Pred. No. 0.24;
 Matches 104; Conservative 74; Mismatches 183; Indels 154; Gaps 22;

```

OY  3  VVASAPKVTMTGTYLLEKPNAGLVSTNARFYAIVKPIKEPKPSMAKMTDVLTS 62
DB  6  LTVASAPKVTLLHGHAAVH-GRVALAVSLNLRFLRLQPSHN-----GKVLSTL 53
OY  63  POLSRESMYKLSLNHLTLOSASDSNRNPFVHAIOYIAAHLATEKDESLKLLQG 122
DB  54  PNIGIKAMQVA-----RLQSLDLS-----FLEQGD-----VTTPESQVQKLR 92
OY  123  LDITLLGSNDPYSYRNOIESAGLP--LTPBSLGTLPAPASTFFMAESNGANSKPREVA- 178
DB  93  -----EVAAGLPDDCAVIERLAIVLA-FLLYLSTICRKRALPSLDIV 133
OY  179  -----KTGLGSSAAMTTAVVAALHLYGVNDSLPCKEG-----KFGCSLDVYHMTAQT 228
DB  134  WSELTPGAGLSSAAYSVCLAAALTY--CEEJPNPLKDGCVNKRKREDELINKWAFQ 191
OY  229  SHCLAGKVGSGEDVSCAVYGSORRYRSEVLSFAQVAVTGLPLNEVIGTILKGMKDNK 288
DB  192  GERIHON-PSQVDNMAVSTWG-----ALRKHQCKTSSLKRSALQILL----- 234
OY  289  RTEFSPLPLMNLLEGEFGSGSSTPSSVAVKWKQMSDEPKARENWONLSDANILETKL 348
DB  235  -TMTKVP-----RMTALVAGVRNRLKPEPEIVAPLTSISALISECERYL 279
OY  349  NDLISKLAKHMDVLYRIRKSCSVLTSEKVLHATEPINEALIKELLEAREMLRIRILMR 408
DB  280  GEMGE-----APAPROYLVLEELIDMNOHHL----- 306
OY  409  OMGEAASVPIEPESOTOLDSTMSABGVLLAGVAGGPAIFAITLTDGSKTLQWASS 468
DB  307  -----ALGVGHASLDQLQVY--RARG-LHSKLTGAGGGC--GITLLKGLDQPEVEAT 356
OY  469  HNVALLIVREDPHGVCLSGSDPRTTCITSGVSIH 503
DB  357  KOALTSCGFD-----CLE-----TSIGARGVSIH 380

```

RESULT 4

```

KINE_PYRHO
ID  KINE_PYRHO  STANDARD;  PRT;  335  AA.
AC  059291;
DT  16-OCT-2001 (Rel. 40, Created)
DT  16-OCT-2001 (Rel. 40, Last sequence update)
DT  16-OCT-2001 (Rel. 40, Last annotation update)
DE  Mevalonate kinase (EC 2.7.1.36) (MK).
GN  MK OR PH1625.
OS  Pyrococcus horikoshii.
OC  Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
OC  Pyrococcus.
OX  NCBI_TaxID=53953;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=OT3;
RX  MEDLINE=9834137; PubMed=9679194;
RA  Kawarabayashi Y., Sawada M., Horikawa H., Haikawa Y., Hino Y.,
RA  Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyama A., Nagai Y.,
RA  Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohnuki Y.,
RA  Funahashi T., Tanaka T., Kudoh Y., Yamazaki J., Kushiida N., Oguchi A.,
RA  Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K.,
RA  Masuchi Y., Shizuya H., Kikuchi H.;
RT  "Complete sequence and gene organization of the genome of a hyper-
RT  thermophilic archaeobacterium, Pyrococcus horikoshii OT3."
RL  DNA Res. 5:55-76(1998).
CC  -I- CATALYTIC ACTIVITY: ATP + (R)-mevalonate = ADP + (R)-5-
CC  phosphomevalonate.
CC  -I- COFACTOR: MAGNESIUM (BY SIMILARITY).
CC  -I- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC  -I- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC  -I- SIMILARITY: BELONGS TO THE GMP KINASE FAMILY. MEVALONATE KINASE
CC  SUBFAMILY:

```

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation
 at the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@sib-sib.ch).

```

CC  EMBL: AP000006; BAA30737.1;
CC  InterPro: IPR001745; GMPkinse_ATP.
CC  InterPro: IPR001459; Mev_gal_kin.
CC  Pfam: PF00288; GMP_kinases; 1.
CC  PRINTS: PR00959; MEVALKINASE.
CC  TIGRFAMs: TIGR00549; mevalon_kin; 1.
CC  DR  PROSITE: PS00627; GMP_KINASES_ATP; 1.
CC  KW  Transferase; Kinase; ATP-binding; Magnesium; Complete proteome.
CC  FT  N_BIND 111 121 ATP (POTENTIAL).
CC  SQ  SEQUENCE 335 AA; 35686 MW; 17D0A9E1D22A1EB1 CRC64;

```

Query Match 4.5%; Score 116; DB 1; Length 335;
 Best Local Similarity 21.9%; Pred. No. 0.22;
 Matches 102; Conservative 60; Mismatches 124; Indels 180; Gaps 25;

```

OY  3  VVASAPKVTMTGTYLLEKPNAGLVSTNARFYAIVKPIKEPKPSMAKMTD-- 57
DB  5  VVASAPKVTLLHGHAAVH-GRVALAVSLNLRFLRLQPSHN-----GKVLSTL 53
OY  58  VKLTSPLSRESMYKLSLNHLTLOSASDSNRNPFVHAIOYIAAHLATEKDESLK 117
DB  62  VSESEDKIETPEYDYGKAAEVL-----YRYATIELAL-----EESIKR--- 99
OY  118  LLLQGLDITLLGSNDPYSYRNOIESAGLP--LTPBSLGTLPAPASTFFMAESNGANSKPREV 177
DB  100  ---VGIDVSI-----TSQIPV-----GA----- 114
OY  178  AKTGIGSSAAMTTAVVAALHLYGVNDSLPCKEGKFGCSLDVYHMTAQTSH---CIAQ 234
DB  115  ---GLSSAANVAVATIGAVSRLLG-LELS-----KE-----EIALGKHVELLVQ 155
OY  235  GKVGSGEDVSCAVYGSORRY--RSEPEVLSFAQVAVTGLPLNEVIGTILKGMKDNK 291

```

```

Db 156 G-ASSGIDPVSAGCEFLYKOGKEP-----LPFME-----186
Qy 292 FSUPLMNLFLGEPGSGSSSTPSVYGAVKKQMSDPKARENNOMSDANILETKINDL 351
Db 187 --LP-----IYVGTGSGT-STKELVAAYRRKRYEMPE-----LVEPILFAMGKLVDK 231
Qy 352 SKLAKDMDVYLVRKSCSVLTSEKVLHATEPINEAIKELLEARERLRIIMR-QM 410
Db 232 AK-----EILSKUDEEKLTKRLGSLMINH 257
Qy 411 GEASVPIEPESOTQLDSTMSAGVLLAGVPGAGFDAIFATLIG 456
Db 258 GLDLDALGVSTKIGELVYAARTA-GAIGAKLTGAGGGGCMYALAPG 302

RESULT 5
KME_MOUSE
ID KIME_MOUSE STANDARD; PRT; 395 AA.
AC 09R08;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE mevalonate kinase (EC 2.7.1.36) (MK).
GN MYK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9930561; PubMed=10401001;
RA Houten S.M., Romeljn G.J., Koster J., Gray R.G.F., Darbyshire P.,
RA Smit G.P.A., de Klerk J.B.C., Duran R., Gibson K.M., Wanders R.J.A.,
RA Waterham H.R.,
RT "Identification and characterization of three novel missense mutations
RT in mevalonate kinase CDNA causing mevalonic aciduria, a disorder of
RT isoprene biosynthesis."
RL Hum. Mol. Genet. 8:1523-1528(1999).
CC -1- FUNCTION: MAY BE A REGULATORY SITE IN CHOLESTEROL BIOSYNTHETIC
CC PATHWAY.
CC -1- CATALYTIC ACTIVITY: ATP + (R)-mevalonate = ADP + (R)-5-
CC phosphomevalonate.
CC -1- ENZYME REGULATION: FARNESYL- AND GERANYL-PYROPHOSPHATES ARE
CC COMPETITIVE INHIBITORS (BY SIMILARITY).
CC -1- PATHWAY: CHOLESTEROL BIOSYNTHESIS, MEVALONATE CATABOLISM.
CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic and peroxisomal.
CC -1- SIMILARITY: BELONGS TO THE GMP KINASE FAMILY. MEVALONATE KINASE
CC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb.sib.ch).
CC -----
DR EMBL: AF137598; AAF00700.1; -.
DR MGD: MGI:107624; MWK.
DR InterPro: IPR0011745; GHMPkinse_ATP.
DR InterPro: IPR0011745; Galkinase.
DR InterPro: IPR001459; mev_gal_kin.
DR Pfam: PF00288; GHMP_kinases.1.
DR PRINTS: PR00960; LMBPPOREIN.
DR PRINTS: PR00959; MEVGALKINASE.
DR TIGRFS: TIGR00549; mevalon_kin.1.
DR TRANSIT: PS00627; GHMP_KINASES_ATP.1.
DR Translase: Kinase; Cholesterol biosynthesis; ATP-binding;
KW Peroxisome.
FT NP_BIND 138 148 ATP (POTENTIAL).
SQ SEQUENCE 395 AA; 41877 MW; 953DB1C89403A3F8 CRC64;

```

```

Query Match 4.4%; Score 113.5; DB 1; Length 395;
Best Local Similarity 21.1%; Pred. No. 0.43;
Matches 96; Conservative 75; Mismatches 152; Indels 131; Gaps 22;

Qy 2 AVVASAGKVLMTGVLVEKPNAGLVLTSTNAREYAIKRPINEEVPKESWAKMTDYKLT 61
Db 5 ALLVASAGKVLHGEHVAHVH-GKVALAALNLPFLRLPQSN-----GKYSVN 52
Qy 62 SPQLSRESMYKLSLNHLTQSVASDSRNPVEYAIQYATAAHLATEKKEKESLHKLLIQ 121
Db 53 LPNIGIQVMDVGM-----LQRLDTS-----FLE---OGDVSVPIL-----EQLEKLRKM 94
Qy 122 GLDITIGSNDVSYRRNOIESAGLPITPESIGTLAPFASI-----TFNAESNGANSKPE 176
Db 95 G-DLP-----RDRAGNEGMA-----LAFILYLAICRKORTLPISLDVMYVSELP 139
Qy 177 VAKTGLSSAAMTTAVVAALLHYGVYDLPDCKEG-----KFGSDLDVYHMAQTSHTL 232
Db 140 GA--GLGSSAAYSVCLAALL--TACEEVSNPDKGVSVSRWPEDELSIKKNAFEGGRV 195
Qy 233 AQGVSGGFVYSCAVYGSORVRESPEVLSFAQVAVGVLPINEYIGTILKCKMKNKTEF 292
Db 196 IHGN-PGVVDNAVSTWGA--LRFQOGTMS---SLKSLPSLIQILTNTK-----238
Qy 293 SLPLMNLFLGEPGSGSSSTPSVYGAVKKQMSDPKARENNOMSDANILETKINDLS 352
Db 239 -VP-----RSTKALVAAYRSRLTKRPEIYAPLTITSDAISLECEVLEGEM- 282
Qy 353 KLANDMDVYLVRKSCSVLTSEKVLHATEPINEAIKELLEARERLRIIMRQMG 412
Db 283 -----VAAPVEQVLYVEELIDMNQHHLN-----306
Qy 413 AASVPIEPESOTQLDSTMSAGVLLAGVPGAGG 446
Db 307 --ALGVGHNSLDLCQYV-AAHG-LHSKLTGAGG 336

RESULT 6
DH1L_RAT
ID DH1L_RAT STANDARD; PRT; 287 AA.
AC P16232;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE corticosteroid 11-beta-dehydrogenase, isozyme 1 (EC 1.1.1.146) (11-DH)
DE (11-beta-hydroxysteroid dehydrogenase 1) (11-beta-HSD1).
GN HSD1B1 OR HSD11.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxId=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=Liver;
RC MEDLINE=90037015; PubMed=2808402;
RA Agarwal A.K., Monder C., Eckstein B., White P.C.;
RT "Cloning and expression of rat cDNA encoding corticosteroid 11 beta-
RT dehydrogenase."
RL J. Biol. Chem. 264:18939-18943(1989).
RN [2]
RP SEQUENCE OF 1-69 FROM N.A.
RX MEDLINE=92375101; PubMed=1508221;
RA Moisan M.P., Edwards C.R., Seckl J.R.;
RT "Differential promoter usage by the rat 11 beta-hydroxysteroid
RT dehydrogenase gene."
RL Mol. Endocrinol. 6:1082-1087(1992).
RN [3]
RP SEQUENCE OF 1-40.
RX MEDLINE=86278718; PubMed=3460996;
RA Monder C., Shackleton C.H.L., Bradlow H.L., New M.I., Stoner E.,
RA Johan F., Laksini V.;
RT "The syndrome of apparent mineralocorticoid excess: its association

```

RT with 11 beta-dehydrogenase and 5 beta-reductase deficiency and some
 RT consequences for corticosteroid metabolism.";
 RT J. Clin. Endocrinol. Metab. 63:550-557(1986).
 RN [4]
 RN MUTAGENESIS OF TYR-179 AND LYS-183.
 RX MEDLINE-93038663; PubMed-1417845;
 RA Obeid J., White P.C.;
 RT "Tyr-179 and Lys-183 are essential for enzymatic activity of 11 beta-
 RT hydroxysteroid dehydrogenase.";
 RT Biochem. Biophys. Res. Commun. 188:222-227(1992).
 CC - FUNCTION: CATALYZES REVERSIBLY THE CONVERSION OF CORTICOSTERONE TO
 CC 11-DEHYDROCORTICOSTERONE.
 CC - CATALYTIC ACTIVITY: An 11-beta-hydroxysteroid + NADP(+) -> an 11-
 CC oxosteroid + NADPH.
 CC - SUBCELLULAR LOCATION: Mitochondrial.
 CC - TISSUE SPECIFICITY: LIVER, KIDNEY, TESTIS AND PLACENTA.
 CC - PTM: GLYCOSYLATED.
 CC - SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES
 CC (SDR) FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>
 CC or send an email to license@sdb.ch).
 CC -----
 DR EMBL: J05107; AAA40886.1; -
 DR EMBL: S43333; AAB22993.1; -
 DR PIR: A34430; DXRTH.
 DR InterPro: IPR002198; ADH_short.
 DR Pfam: PF00106; adh_short.1.
 DR PROSITE: PS00061; ADH_SHORT.1.
 KW Oxidoreductase; Glycoprotein; NADP; Microsome; Steroid metabolism.
 FT NP_BIND 30 78 NADP (BY SIMILARITY).
 FT ACT_SITE 179 179 D->N: SLIGHT LOSS OF ACTIVITY.
 FT MUTAGEN 110 110 Y->F,S: COMPLETE LOSS OF ACTIVITY.
 FT MUTAGEN 179 179 Y->F,S: COMPLETE LOSS OF ACTIVITY.
 FT MUTAGEN 183 183 K->R: COMPLETE LOSS OF ACTIVITY.
 SQ SEQUENCE 287 AA; 31769 MW; 6E532CFA9ACF1E22 CRC64;
 Query Match 4.3%; Score 112; DB 1; Length 287;
 Best Local Similarity 24.6%; Pred. No. 0.35;
 Matches 52; Conservative 27; Mismatches 82; Indels 50; Gaps 8;
 QY 43 NEEVRESNAWKTVKLTSPOLRESMYKLSL--NHLTLOSASDSRNPFEVHAIOYA 100
 DB 20 NEEFREMLQGGKVIYTGSKGIGREMAHILSKMGAVHL-TARSEEGLOKYSRCLIEG 78
 QY 101 IAAAH--LATEKDESLHKL-----LLOGLDITTLG-----SNDFSYRNOIE- 141
 DB 79 AASAHYIAGTMEDEMAFEFVVEAGKLGIDMLILNHLITQTMISLFHDIHSRSMVEV 138
 QY 142 -----SAGPLRPESLGTLPAPASITFNAAESNGANSKREPAKGTGGSSA----- 186
 DB 139 NELSIVLSTAAALPMLKOSNGSIAIISMA-----GKMTPLASYSASKRFALDGFES 191
 QY 187 -----AMTAAVVAALILHYLGVDLSDPCKE 211
 DB 192 TIRKEHLMTKVNSITLVCVGFIDTETALKE 222
 RESULT 7
 KIME PYRAB
 ID KIME PYRAB STANDARD: PRT; 335 AA.
 AC Q9V187;
 DT 16-OCT-2001 (rel. 40; Created)
 DT 16-OCT-2001 (rel. 40; Last sequence update)
 DT 16-OCT-2001 (rel. 40; Last annotation update)
 DE Mevalonate kinase (EC 2.7.1.36) (MK).
 GN MK OR PAB0372.
 OS Pyrococcus abyssi.

OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
 OC Pyrococcus.
 OX NCBI_TaxID=29292;
 RN [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=GES / Orsay;
 RA Heilig R.;
 RT "Pyrococcus abyssi genome sequence: insights into archaeal chromosome
 RT structure and evolution.";
 RT Submitted (Jul-1999) to the EMBL/GenBank/DBJ databases.
 CC - CATALYTIC ACTIVITY: ATP + (R)-mevalonate -> ADP + (R)-5-
 CC phosphomevalonate.
 CC - COFACTOR: MAGNESIUM (BY SIMILARITY).
 CC - SUBUNIT: HOMODIMER (BY SIMILARITY).
 CC - SUBCELLULAR LOCATION: Cytoplasmic (Probable).
 CC - SIMILARITY: BELONGS TO THE GMP KINASE FAMILY. MEVALONATE KINASE
 CC SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>
 CC or send an email to license@sdb.ch).
 CC -----
 DR EMBL: AJ248284; CAB49463.1; -
 DR InterPro: IPR001745; GMPkinase_ATP.
 DR InterPro: IPR001174; GALKkinase.
 DR InterPro: IPR001459; Mv_gal_kin.
 DR Pfam: PF00288; GMP_Kinases.1.
 DR PRINTS: PR00960; LMPPROTEIN.
 DR PRINTS: PR00959; MEVGAALKINASE.
 DR TIGRPS: TIGR00549; mevalon_kin.1.
 DR PROSITE: PS00627; GMP_KINASES_ATP.1.
 KW Transferase; Kinase; ATP-binding; Magnesium; Complete proteome.
 FT NP_BIND 111 121 ATP (POTENTIAL).
 SQ SEQUENCE 335 AA; 35774 MW; EDB06EDA186599C CRC64;
 Query Match 4.3%; Score 112; DB 1; Length 335;
 Best Local Similarity 20.2%; Pred. No. 0.44;
 Matches 102; Conservative 67; Mismatches 139; Indels 198; Gaps 24;
 QY 3 VVASAPGKVLMTGCV-LVLEKPNAGLVLTNARFVIAVKPNE--EVKPESSNAWKTD-- 57
 DB 5 VLASAPAKIILFEHGSVYVQKP--AIASAIDLFTY-VRAEFNSGNIKIEAHDKIPGLI 61
 QY 58 VKLTSPOLRESMYKLSLNLHLTLOSASDSRNPFEVHAIOYAAAHLATEKDESLK 117
 DB 62 VSESEDKITFETDYGRAAEVLS-----YVRAHLEVL-----EADKRT--- 100
 QY 118 LLLQGLDITILGSDNFYSYRNQIESAGLPLTPESLGTLPAPASITFNAAESNGANSKREPV 177
 DB 101 -----GVSVSI-----TSQIPV-----GA----- 114
 QY 178 AKTGGLSSAAMTAAVVAALILHYLGVDLSDPCKEKGFGSDLVLT-HMTAQTH---CLA 233
 DB 115 -----GLSSAANAATAVATIGAVSKL-----DLEISKEELAKMGHVELLY 154
 QY 234 QGKVGSGFDVSCAVGSGQRYVRSPVLSPAQAVATGVLPLNEVIGTILGKMDKRTRES 293
 DB 155 QG-ASSGIDPTYSATIGFLY-----KQGEFE 180
 QY 294 LPTIMNL--FLGPPGSGGSSPTSMGAVKKMQMSDEKAREMNQNLSDANLELETKLNL 351
 DB 181 HLPFVELPIYVGTGSSG-STKELVAMVRRRYEME-----LIEPLIEBMSGLTVK 231
 QY 352 SKLAKHMDVLYLRVKSIVLSEKVVLAHTEPINDAIIKELLEAREAMRITILMR-QM 410
 DB 232 AK-----EVIISK-LDEBEKEFKLGELENNINH 257
 QY 411 GEASVPIEPESOTQLDSTMSAEGVLLACVPPAGGFDATFATLTDSDSGKYLQAMSSNH 470

Db 258 GLLDALGVSTKKLSLYAANTA-GAIGAKITGAGGGCMYALAPGK----- 304

Qy 471 VLALVREDPHGVCSGSDPRTTCT 496

Db 305 -----REVATAIKIAGTPTMTTIS 324

RESULT 8

TIME RAT STANDARD; PRT; 395 AA.

AC p17256;

DT 01-AUG-1990 (Rel. 15, Created)

DT 01-AUG-1990 (Rel. 15, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Mevalonate kinase (EC 2.7.1.36) (MK).

GN MYK.

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

OX NCBI_TaxID=101116;

RN [1]

RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-20.

RC STRAIN=Sorogue-Dawley; TISSUE=Liver;

RA MEDLINE=90222132; PubMed=2138094;

RA Tanaka R.D., Lee L.Y., Schafer B.L., Kratunis V.J., Mohler W.A., Robinson G.W., Mosley S.T.;

RT "Molecular cloning of mevalonate kinase and regulation of its mRNA levels in rat liver."

RT Proc. Natl. Acad. Sci. U.S.A. 87:2872-2876(1990).

CC -1- FUNCTION: MAY BE A REGULATORY SITE IN CHOLESTEROL BIOSYNTHETIC PATHWAY.

CC -1- CATALYTIC ACTIVITY: ATP + (R)-mevalonate = ADP + (R)-5-phosphomevalonate.

CC -1- ENZYME REGULATION: FARNESYL- AND GERANYL-PYROPHOSPHATES ARE COMPETITIVE INHIBITORS.

CC -1- PATHWAY: CHOLESTEROL BIOSYNTHESIS, MEVALONATE CATABOLISM.

CC -1- SUBUNIT: HOMODIMER.

CC -1- SUBCELLULAR LOCATION: Cytoplasmic and peroxisomal.

CC -1- DISEASE: MUTATION IN THE MEVALONATE KINASE GENE CAUSES MEVALONIC ACIDURIA.

CC -1- SIMILARITY: BELONGS TO THE GMP KINASE FAMILY. MEVALONATE KINASE SUBFAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL: M29472; AAA41588.1; -

DR PIR: A35629; A35629.

DR InterPro: IPR001745; GMPKase-ATP.

DR InterPro: IPR001459; Mev_gal_kin.

DR Pfam: PF00288; GMP_kinases; 1.

DR PRINTS: PRO0959; MEYGALKINASE.

DR TIGRFAMs: TIGR00549; mevalon_kin; 1.

DR PROSITE: PS00627; GMP_KINASES-ATP; 1.

DR Translasease: Kinase; Cholesterol biosynthesis; ATP-binding; Peroxisome.

KW NP_BIND 138 148 ATP (POTENTIAL).

FT SEQUENCE 395 AA; 41987 MM; 803DIF44E3C525FC CRC64;

SO

Query Match 4.38; Score 111.5; DB 1; Length 395;

Best Local Similarity 20.98; Pred. No. 0.61;

Matches 98; Conservative 67; Mismatches 142; Indels 161; Gaps 22.

Qy 3 VVAAPGKVLMTGGLVLEKPNAGVLTSTNARFAIVKPIKEEVKPESMAMKWDVKTLS 62

Db 6 LTVAPGKVLMTGGLVLEKPNAGVLTSTNARFAIVKPIKEEVKPESMAMKWDVKTLS 62

Qy 63 POLSRESMYKLSLNLHLLTLOSVSASDSRNPVEVHAIOYVAIAAAHLATEKDESLHKLILQG 122

Db 54 PNVGKQWADVA-----TLQLDPTG-----FLE---QGDVPAPTL-----EQLEKTK--- 92

Qy 123 LDITILGSDNDFSYNQLSAGLP---LTPSELGTLA-----PFASITFN 164

Db 93 -----KVAGLPDPCVGNBSGLSLATLFLYLALICRQRLLPSLDIMW 134

Qy 165 AAESNGANSKEVATGTLGSSAAMTTAVVAALLHYLGVVDSLPCKE---GRFGCSDDL 220

Db 135 SELPFGA-----GLGSSAAYSVCVAALL--TACEEVTNPLKDGSGISGWEEELK 183

Qy 221 VIHMTAQTSHCLAGKVGSGPDVSCAVYGSQRYRFSPEVLVSFAOAVATGPLENVTGT 280

Db 184 SINKAAVGEERYIHGN-PSGVDNSVSWGCA--LRYOOGKMS---SLKRLPALQITLTN 236

Qy 281 LKGKMDNRTESLPPLNLLFLGEGSGSSPPSVGAVKKQMSDPKARENMQLSDA 340

Db 237 TK-----VP-----RSTKALVAGVRSLIKFPEIMAPLTSIDAI 271

Qy 341 NLEETKINDSLKAKDMDVLYRIVKCSVLTSEKWLHATEPINE-AIKELLEARE 398

Db 272 SLECCERVIGEM-----AAAPVPEQVILELMDNQ 302

Qy 399 AMLRIRIIMRONGEASVPTEPESQTLIDTMSAEGVVLACVPGAG 446

Db 303 HHLN-----ALGVGHASLDQLCOYT-AAHG-LHSKLTGAGG 336

RESULT 9

DH11_MOUSE STANDARD; PRT; 292 AA.

ID DH11_MOUSE

AC P50172;

DT 01-OCT-1996 (Rel. 34, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Corticosteroid 11-beta-dehydrogenase, Isozyme 1 (EC 1.1.1.146) (11-DH) (11-beta-hydroxysteroid dehydrogenase 1) (11-beta-HSD1) (11beta-HSD1A).

GN HSD11A1 OR HSD11.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6 X CBA; TISSUE=Liver;

RA MEDLINE=95178382; PubMed=7873449;

RA Rajan V., Chapman K.E., Lyons V., Jamieson P., Mullins J.J., Edwards C.R., Seckl J.R.;

RT "Cloning, sequencing and tissue distribution of mouse 11 beta-hydroxysteroid dehydrogenase-1 cDNA."

RT Eur. J. Biochem. 227:202-208(1995).

RL J. Steroid Biochem. Mol. Biol. 52:141-147(1995).

RN [2]

RN [3]

RP SEQUENCE FROM N.A.

RC STRAIN=BALB/c; TISSUE=Liver;

RA MEDLINE=95154291; PubMed=7851387;

RA Opperman U.C.T., Netter K.J., Maser E.;

RT "Cloning and primary structure of murine 11 beta-hydroxysteroid dehydrogenase/microsomal carbonyl reductase."

RT Eur. J. Biochem. 227:202-208(1995).

RL [3]

RP SEQUENCE OF 1-10 FROM N.A.

RA Voice M.W.;

RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.

CC -1- FUNCTION: CATALYZES REVERSIBLE THE CONVERSION OF CORTISOL TO THE INACTIVE METABOLITE CORTISONE.

CC -1- CATALYTIC ACTIVITY: An 11-beta-hydroxysteroid + NADP(+) = an 11-oxosteroid + NADPH.

CC -1- SUBCELLULAR LOCATION: Microsomal.

CC -1- TISSUE SPECIFICITY: WIDELY EXPRESSED. HIGHEST EXPRESSION IN LIVER.

CC -1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES (SDR) FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: S75207; AAB3601.1; -
 DR EMBL: X83202; CA58209.1; -
 DR EMBL: X92186; CA63096.1; -
 DR MGD: MG1:103562; Hsd1bl
 DR InterPro: IPR002198; ADH_short.
 DR Pfam: PF00106; adh_short; 1.
 DR PROSITE: PS00061; ADH_SHORT; 1.
 DR Oxidoreductase; NADP; Microsome; Steroid metabolism.
 FT NP_BIND 34 63 NADP (BY SIMILARITY).
 FT ACT_SITE 183 183 BY SIMILARITY.
 FT CONFLICT 15 15 F -> S (IN REF. 2).
 FT CONFLICT 232 232 N -> D (IN REF. 2).
 FT CONFLICT 234 234 Q -> L (IN REF. 2).
 FT CONFLICT 261 261 S -> L (IN REF. 2).
 SQ SEQUENCE 292 AA; 32364 MW; ADE42B1D82DD6CD CRC64;

Query Match 4.28; Score 109.5; DB 1; Length 292;
 Best Local Similarity 21.7%; Pred. No. 0.55;
 Matches 70; Conservative 39; Mismatches 109; Indels 105; Gaps 14;

QY 27 LVISTNARFAIVKPIEEFKPESMAKMTDVKLTSPOLRESMYKSL--NHLTQSVS 84
 DB 12 LVFLAYIYS-----TNEEFREMLQCKKVIYVGASGIGREMAVHLISKGAHVYL-TAR 66
 QY 85 ASDSRNPFVEHAIOYAIYAAH--LATEKRESLHKL-----LLQGLDITILG----- 129
 DB 67 SEEGIKVRSRCLELGAASHAYIAGTMEQTFAPQFIVKAGKLGMDLNLNHIQTSL 126
 QY 130 ---SNDYSYRNQIE-----SAGLPTEPSLGTLPASFITNNAESNANKSKE 176
 DB 127 SLFHDHDSVRVWEVNFSLSYVNSTALPLKOSNGSINAVISLA-----GKMPQPM 179
 QY 177 VAKTGLSSAA-----MTTAVVALLHYLGVNLDSPCKEKGCCSDLDVTHM 224
 DB 180 IAVYSASKFLDGFSTIRRELYTKKNSVITLCVGLIDETAMKE----- 226
 QY 225 IADTSHCLAQKVGSGFDVSCAVYSGQRYVRSPEVLSPAQAVATGPLEVI--GTILKG 283
 DB 227 ISGIINQASPK-----EECAL-----EIIQTALR- 252
 QY 284 KMDNKRTEPSLPLMLFLGEPG 306
 DB 253 ---KSEVYDYKSPLETPILLGNPG 272

RESULT 10
 FTSH_CAPAN STANDARD; PRT; 662 AA.
 AC Q39444;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Cell division protein ftsh homolog, chloroplast precursor
 DE (EC 3.4.24.-) (Fragment).
 GN FTSH.
 OS Capsicum annuum (Bell pepper).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; eustersids I; Solanales; Solanaceae; Capsicum.
 NCBI_TaxID=4072;
 RN (1)
 RP SEQUENCE FROM N.A.
 RA Houine G., Schantz M.L., Schantz R.;
 RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
 CC -i- FUNCTION: SEEMS TO ACT AS AN ATP-DEPENDENT ZINC METALLOPEPTIDASE

CC (BY SIMILARITY).
 CC -i- COFACTOR: BINDS 1 ZINC ION (POTENTIAL).
 CC -i- SUBCELLULAR LOCATION: CHLOROPLAST; INTEGRAL MEMBRANE PROTEIN
 CC (POTENTIAL).
 CC -i- SIMILARITY: BELONGS TO THE AAA FAMILY OF ATPASES.
 CC -i- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M41.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: X90472; CA62084.1; -
 DR MEROPS: M41.005; -
 DR InterPro: IPR003593; AAA_ATPase.
 DR InterPro: IPR003959; AAA_ATPase-centr.
 DR InterPro: IPR003960; AAA_sub.
 DR InterPro: IPR000642; Peptidase_M41.
 DR Pfam: PF00004; AAA; 1.
 DR Pfam: PF01434; Peptidase_M41; 1.
 DR SMART: SM00382; AAA; 1.
 DR TIGRfam: TIGR01241; FtsH_fam; 1.
 DR PROSITE: PS00674; AAA; 1.
 KW Cell division; ATP-binding; Transmembrane; Hydrolase; Metalloprotease;
 KW Zinc; Chloroplast; Transil peptide.

FT TRANSIT 1 1 CHLOROPLAST (POTENTIAL).
 FT NON_TER 1 1
 FT CHAIN 1 662 CELL DIVISION PROTEIN FTSH HOMOLOG.
 FT TRANSMEM 35 35 POTENTIAL.
 FT TRANSMEM 173 193 POTENTIAL.
 FT NP_BIND 270 277 ATP (POTENTIAL).
 FT METAL 492 492 ZINC (CATALYTIC) (BY SIMILARITY).
 FT ACT_SITE 493 493 BY SIMILARITY.
 FT METAL 496 496 ZINC (CATALYTIC) (BY SIMILARITY).
 FT NON_TER 662 662
 SQ SEQUENCE 662 AA; 71061 MW; AF3411BDC502C6A1 CRC64;

Query Match 4.28; Score 109; DB 1; Length 662;
 Best Local Similarity 20.8%; Pred. No. 2;
 Matches 119; Conservative 75; Mismatches 209; Indels 168; Gaps 28;

QY 2 AVVASAPGKVLMTGVLVEKPNAG---LVISTNARFYA---IVKPINEEKPSMAK 54
 DB 55 AVIASPP-----QVMEVEAPNPNTSNPLPEQNLVLPKQASPSVD--LPESQWR 105
 QY 55 WTDVKLTSPOLSRSMKSLNHLTLOSVAASDSRNFVEHAIOYAIYAAH-LATEKRES 114
 DB 106 YSEF-LNAVKKKVERVRFSSKDGSAIQ-LTAVDGR-----NAVIVPNPDDL 150
 QY 115 LHKLLQGLDITIL---GSDNFYSYRNQIESAGLPTEPSLGTLPASFITNNAESNGA 171
 DB 151 IDILAMGVDSIVSEGGNGLEFSVGNL-----LFPPIAAGLFFLFRSQGG 199
 QY 172 NSKP-----EVAKTGLGSSAAMTTAVVA---ALLHYLGVND-LSDPC 209
 DB 200 PGCGGGLGPMDFGRSKSKQVEPTEG---VTFADVAGADQAKLELEVDVFLNPD 253
 QY 210 KEKFGCSDLDVHMIADTSHCLAQKVGSGFDV-----SCAY----- 247
 DB 254 KYTALGAK-----IPKCLLVGPPGKGTLLARAVAGEAGVPPFSCAASEVELFV 304
 QY 248 -YSGQR-----VRSPEVLSFAQVAVTGL-----PLNEVIGTILK 284
 DB 305 GVGASRVRLHLENAKSKAPCVIFIDEIDAVGRQAGAGGNDEREOTINQLL-TENDGF 363
 QY 285 WDNK-----RTEFSLPLMLFLGEGSGGSST-----PSMYGAAYKKMQMSPDKA----- 330
 DB 364 SGNGSVLYLATATNRPVDLSALLRPGKFDQRYVYDRDVGARVAILLOYHSGRLAADVD 423
 QY 331 -----RENWNLSDANLETFKLNDLSKLAKDH-WDYVLRVI-----KSCSVLT 373

Db 424 FDKIARTPGFTGADLQINMEAAIILAR-RDLKEISKDEISDALERTIACPEKKNAVVS 482
Oy 374 SEKVLAHTEPINEAIKELLEAREAMRILROROEASVPEPSQOL---LDST 430
Db 483 DEKKLVAHAGHALVAGALPEYDPAVKISIPR--GQAGLTFEPASERLESGLYSR 540
Oy 431 MSAGVLLAGVAGGAFDAFA---ITLGS 458
Db 541 SYLENOMAVALGGRVAREVITGEDNVTGAS 571
RESULT 11
NXIA MOUSE STANDARD: PRT: 529 AA.
AC 09CS84: 088722;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE Neurexin 1-alpha (Neurexin I-alpha) (Fragments).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus;
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE OF 1-140 FROM N.A.
RC STRAIN-CD-1: TISSUE-Brain;
RA Graveley B.R., Phillips D.L.;
RT "Sequencing of the neurexin genes."
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 141-529 FROM N.A.
RC STRAIN-C57BL/6J: TISSUE-Embryo;
RX MEDLINE-21085660; PubMed-11217851;
RA Kawai J., Shimagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,
RA Alizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadoya K., Matsuda H.A., Ashburner M., Batilov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochwa H.,
RA Knehl P., Lewis S., Matsuo Y., Nakado I., Pesole G., Quackenbush J.,
RA Schirral L.M., Staabli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojuna N., Carinci P., de Bernaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya K., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohlski S.,
RA Hayashizaki Y.,
RT "Functional annotation of a full-length mouse cDNA collection."
RL Nature 409:685-690(2001).
RN [3]
RP SEQUENCE OF 478-520 FROM N.A.
RC STRAIN-C57BL/10: TISSUE-Brain;
RA Gorecki D.C., Sklarczyk A., Lukasik K., Kaczmarek L., Simons J.P.,
RT "Differential seizure-induced and developmental changes of neurexin
expression."
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: NEURONAL CELL SURFACE PROTEIN THAT MAY BE INVOLVED IN
CELL RECOGNITION AND CELL ADHESION. MAY MEDIATE INTRACELLULAR
SIGNALING.
CC -1- SUBUNIT: THE CYTOPLASMIC C-TERMINAL REGION BINDS TO CASK (BY
SIMILARITY).
CC -1- SUBCELLULAR LOCATION: TYPE I membrane protein (Potential).
CC -1- ALTERNATIVE PRODUCTS: MULTIPLE ISOFORMS ARE PRODUCED BY
ALTERNATIVE SPLICING (BY SIMILARITY).
CC -1- SIMILARITY: CONTAINS AT LEAST 2 LAMININ G-LIKE DOMAINS.
CC -1- SIMILARITY: BELONGS TO THE NEUREXIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: AF387674; AAK70469.1; -
DR EMBL: AF387674; AAK70470.1; -
DR EMBL: AF387674; AAK70471.1; -
DR EMBL: AK017578; BAB30815.1; -
DR EMBL: AJ006802; CAA07257.1; -
DR HSSP: 063373; 1C4R.
DR MGD: MGI:1096391; Nrxn1.
DR InterPro: IPR001791; Laminin_G.
DR Pfam: PF00054; Laminin_G; 1.
DR SMART: SM00282; Laminin_G; 1.
DR PROSITE: PS50025; LAM_G_DOMAIN; 1.
KW Transmembrane; Cell adhesion; Repeat; Glycoprotein;
KW Alternative splicing.
FT NON_TER 1
FT NON_CONS 140 141
FT DOMAIN <1 453
FT TRANSMEM 454 474
FT DOMAIN 475 529
FT DOMAIN 15 140
FT DOMAIN 178 333
FT CARBOHYD 245 245
FT VARSPLIC 90 96
FT VARSPLIC 82 96
SQ SEQUENCE 529 AA; 57193 MW; 282E04A03DEBA671 CRC64;
Query Match 4.18; Score 107; DB 1; Length 529;
Best Local Similarity 23.48; Pred. No. 2;
Matches 71; Conservative 41; Mismatches 112; Indels 80; Gaps 16;
Oy 9 GKVLITG---GYLVKEKNAGLIVSTN---AREVAIVKPIKEVEPESMAVKWTFDKLTS 62
Db 22 GLMLHTGASADYVNLALNKAGVSLVINGSCAFELVPEVCKRFNDNA---WEDVKT- 76
Oy 63 POLRESMYKSLMLNLTLOSASDSRNPFEHAIOVAIAAHLATEKESLHKLILG 122
Db 77 ---NLNQHSGSIGHAMVKNLHCS-----VTISVDGLITVGTQGE----- 113
Oy 123 LDITLGSNDYSTRNQLSAGLPTP--ESLGTIAPR-----STFNAESNGANSKPE 176
Db 114 -DYTMGLGSDDFEYVGSSTADLPSPVCNDPGTYIFSKGGQITTYWMPNDPSTRAD 172
Oy 177 VAKTGLGSSAMTTPAVNALHLYLGVLDLSPCKEGKFGCSLDVYIHMIAGTSHCLAQK 236
Db 173 --RLAIGSTYQKEAVLVR-----VDS--GLGDYLEH-----IHGK 208
Oy 237 VSGGFV-----SCAVYGSORY--VRESPEVLSFAOVAVTGLPINEYIGTLKRW 285
Db 209 IGVKNVGTGDIALIESNAIINDGKYHVRFTRSG--GNATLOVDSP---VIERYPACNN 264
Oy 286 DNKR 289
Db 265 DNER 268
RESULT 12
Y99F-YEAST STANDARD: PRT: 1442 AA.
AC P47169;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical 161.2 kDa protein in NMD5-HOM6 intergenic region.
GN Y0R137C OR J2126.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;


```

OY 321 KWONS-----DEKARENMONISDANLEETKINDLSKLAKDHMDVLYRYK 367
DB 426 ECGSGYTDINLLGRYNPER-----YDWMSEK--EIESAINLND--AFENGD--RIPE 475
OY 368 SCASYLSEKWLHATEPINEAIKELLEAREAMLRILRQMGAAASVPI-----418
DB 476 GITYVRAOS--MTAFIYALVKNKV-----FYRNFVSTSLPIIFGRFGITH 521
OY 419 -----EPESOTOL-IDSTMSAEGLV 438
DB 522 AGIGLEPEARNELTVDK--NEEGITI 546

RESULT 14
KIME_ARCFU
ID KIME_ARCFU STANDARD; PRT; 284 AA.
AC 027985;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE MEVALONATE KINASE (EC 2.7.1.36) (MK).
GN MKV OR AF2289.
OS Archaeoglobus fulgidus.
OC Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;
OX Archaeoglobaceae; Archaeoglobus.
RN NCBI_TaxID=2234;
RP SEQUENCE FROM N.A.
RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
RX MEDLINE=98049343; PubMed=9389475;
RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
RA Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,
RA Richardson D.L., Kerlavage A.R., Graham D.E., Kyrtides N.C.,
RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
RA Peterson S., Ketch C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,
RA Cotton M.D., Spriggs T., Arlanch P., Kaine B.P., Sykes S.M.,
RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA Venter J.C.;
RT "The complete genome sequence of the hyperthermophilic, sulphate-
RT reducing archaeon Archaeoglobus fulgidus."
RL Nature 390:364-370(1997).
CC -1- CATALYTIC ACTIVITY: ATP + (R)-mevalonate -> ADP + (R)-5-
CC phosphomevalonate.
CC -1- COFACTOR: MAGNESIUM (BY SIMILARITY).
CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -1- SIMILARITY: BELONGS TO THE GMP KINASE FAMILY. MEVALONATE KINASE
CC SUBFAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AE000946; AAB8965.1;
CC TIGR: AF2289;
CC InterPro: IPR001745; GMPkinase_ATP.
CC InterPro: IPR001459; Mev_gal_kin.
CC Pfam: PF00288; GMP_kinases_1.
CC PRINTS: PR00959; MEVGALKINASE.
CC TRIGRAMS: TIGR00549; mevalon_kin_1.
CC PROSITE: PS00627; GMP_KINASES_ATP; 1.
CC TRANSFERASE; Kinase; ATP-binding; Magnesium; Complete proteome.
CC NP_BIND 86 ATP (POTENTIAL).
CC SEQUENCE 284 AA; 30877 MW; F102C1C71772CA7F CRC64;
Query Match 4.0%; Score 103.5; DB 1; Length 284;

```

```

Best Local Similarity 18.4%; Pred. No. 1.5;
Matches 90; Conservative 61; Mismatches 126; Indels 213; Gaps 20;

OY 3 VVASAPGKVLMTGGYVLLEKPNAGLVLTSTNARFAIYKPIKEVYKPPESMAKMDVLTLS 62
DB 1 MIASAPGKILLEGHAYVGR-----HAAVSAINLRCR-----VSXKSD 40
OY 63 POLRESMYKLSLNLHFLQSVASDSRNPEVHAIOVAIAAHLATEKDEKSLKLLLOG 122
DB 41 RFLRSLSGESGLDY-----QRHPYVQAVK----- 66
OY 123 LDITILGNDFFSYRNOIESAGLPLTPESLGTLPASPSTFMAESNGANSKPREVANTGL 182
DB 67 -----RFGELRN-IPGAEIEESE-----IPIGS-----GL 91
OY 183 GSSAAMTAVVAALHLHGLVVDLSDPCKEKF--GCSLDYIHMTAQTSHCLAQKVGSGF 241
DB 92 GSSAAVIVATIAL-------NAEFDGDMKEALFQAKOVEIDVGR--ASGI 136
OY 242 DVSCAVVGSQRVYFSEVLFSFAOVAVTGLPLNEVIGITLKGW---DNKRTEFSLPLM 298
DB 137 DPFISTFG-----GSMLFPERKVEMPKFFV 163
OY 299 NLFGEPSGGSSTPSNVGAVKWKWMSDPEKARENMONISDANLEETKINDLSKLAKDH 358
DB 164 INF-----GSRSTAEMVAKVAELRERHPEVDKIFDAIDALSIE-----202
OY 359 MDVLYRYKSCSVLTSEKWLHATEPINEAIKELLEAREAMLRILRQMGAAASVPI 418
DB 203 -----ASDVSAER-----LEELIALINOSLIR-----AIGVS- 229
OY 419 EPESOTOLDSTMS--AEGVLLAGVPGAGPDFAFATILGSGSTRKIQAWSHNVALLV 476
DB 230 NPE-----IDTIALLELMGMNAKITGAGGCGCFLGFKGKPRG-----SRTV 273
OY 477 REDPHVCYLE 486
DB 274 EPEKEGVATIE 283

RESULT 15
LDHA_BRARE
ID LDHA_BRARE STANDARD; PRT; 332 AA.
AC 09PVK5;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE L-lactate dehydrogenase A chain (EC 1.1.1.27) (LDH-A).
GN LDHA.
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7953;
RN 11
RP SEQUENCE FROM N.A.
RA Tsou S.C.-M., Li J.Y., Mannen H., Li S.S.-L.;
RT "Molecular evolution of vertebrate lactate dehydrogenase isozymes by
RT gene duplication."
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: (S)-lactate + NAD(+) -> pyruvate + NADH.
CC -1- PATHWAY: Anaerobic glycolysis; final step.
CC -1- SUBUNIT: Homotrimer (By similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: BELONGS TO THE LDH FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```


GenCore version 5.1.4.p5.4578.
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 26, 2003, 12:50:30 ; Search time 39 Seconds
(without alignments)
1725.425 Million cell updates/sec

Title: US-09-988-863A-2

Perfect score: 2581

Sequence: 1 MAVVASAPGKMTMGTYLV.....ESGDPRTTCISGVSTHLE 505

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

A.GeneSeq_101002:*

1: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1980.DAT:*

2: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1981.DAT:*

3: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1982.DAT:*

4: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1983.DAT:*

5: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1984.DAT:*

6: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1985.DAT:*

7: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1986.DAT:*

8: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1987.DAT:*

9: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1988.DAT:*

10: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1989.DAT:*

11: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1990.DAT:*

12: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1991.DAT:*

13: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1992.DAT:*

14: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1993.DAT:*

15: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1994.DAT:*

16: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1995.DAT:*

17: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1996.DAT:*

18: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1997.DAT:*

19: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1998.DAT:*

20: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA1999.DAT:*

21: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA2000.DAT:*

22: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA2001.DAT:*

23: /SID52/gcgdata/geneSeq/geneSeq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	2581	100.0	505	23 ABB77505	Arabiopsis thaliana
2	480	18.6	432	22 AAU15094	Protein encoded by
3	475	18.4	432	22 AAY72679	Candida albicans p
4	372.5	14.4	109	21 AAB18130	Pinus radiata phos
5	178.5	6.9	368	22 AAU35063	Enterococcus faeca
6	178.5	6.9	368	22 AAB60859	Mevlonate pathway
7	155	6.0	374	22 AAB99728	Streptomyces sp. C
8	149.5	5.8	359	23 ABB48617	Listeria monocytog
9	147.5	5.7	361	22 AAB60881	Mevlonate pathway
10	142.5	5.5	358	22 AAG81912	S. epidermidis ope

11	142.5	5.5	358	22 AAB60883	Mevlonate pathway
12	142.5	5.5	362	23 ABB40510	Staphylococcus epi
13	136.5	5.3	358	22 AAB60864	Mevlonate pathway
14	135.5	5.2	358	22 AAU37273	Staphylococcus aur
15	133.5	5.2	345	22 AAU34358	Staphylococcus aur
16	130.5	5.1	358	22 AAB60882	Mevlonate pathway
17	123	4.8	378	21 AAG12901	Arabiopsis thaliana
18	123	4.8	378	21 AAG50090	Arabiopsis thaliana
19	123	4.8	378	21 AAG51424	Arabiopsis thaliana
20	123	4.8	380	21 AAG51423	Arabiopsis thaliana
21	123	4.8	398	21 AAG50089	Arabiopsis thaliana
22	123	4.8	447	21 AAG50074	Arabiopsis thaliana
23	123	4.8	451	21 AAG50073	Arabiopsis thaliana
24	123	4.8	456	21 AAG50072	Arabiopsis thaliana
25	117	4.5	396	12 AAR13720	Human mevalonate k
26	116.5	4.5	952	22 AAM78089	Human protein SEQ
27	115.5	4.5	723	22 AAM80073	Human protein SEQ
28	115	4.5	432	21 AAY43633	Amino acid sequenc
29	115	4.5	816	23 ABB53952	Lactococcus lactis
30	114	4.4	823	22 ABB71484	Drosophila melanog
31	112	4.3	287	23 AAO14408	Rat corticosteroid
32	112	4.3	338	22 AAB96178	Putative P. abyssi
33	109.5	4.2	369	11 AAR05738	glyceroldehyde-3-p
34	108.5	4.2	1477	14 AAR41728	High molecular wei
35	108.5	4.2	1477	15 AAR63506	Haemophilus high m
36	108.5	4.2	1477	21 AAB01848	Haemophilus influe
37	108	4.2	492	22 ABB06982	Novel human diagno
38	108	4.2	492	22 ABB15375	Novel human diagno
39	106	4.1	1056	21 AAG31894	Arabiopsis thaliana
40	106	4.1	1073	21 AAG31893	Arabiopsis thaliana
41	106	4.1	1087	21 AAG31892	Arabiopsis thaliana
42	105.5	4.1	308	22 AAM79592	Human protein SEQ
43	105.5	4.1	931	22 AAM78649	Human protein SEQ
44	105.5	4.1	949	22 AAB12315	Human protein SEQ
45	105.5	4.1	949	22 AAM79633	Human protein SEQ

ALIGNMENTS

RESULT 1	ABBT7505	standard; Protein; 505 AA.
ID	ABBT7505	
XX	ABBT7505;	
AC	ABBT7505;	
XX	29-JUL-2002 (first entry)	
DT	29-JUL-2002	
DE	Arabiopsis thaliana PMVK SEQ ID NO 2.	
XX	Thale cress: PMVK, phosphomevalonate kinase; plant; herbicide;	
KW	growth regulator; enzyme.	
KW	Arabiopsis thaliana.	
OS	DE10057755-AL.	
XX	23-MAY-2002.	
PD	22-NOV-2000; 2000DE-1057755.	
XX	22-NOV-2000; 2000DE-1057755.	
PF	22-NOV-2000; 2000DE-1057755.	
XX	22-NOV-2000; 2000DE-1057755.	
PR	(FARB) BAYER AG.	
PA	Weissner R, Lechelt-Kunze C;	
XX	WPI; 2002-445360/48.	
XX	N-PSDB; ABL60244.	
DR	New nucleic acid encoding plant phosphomevalonate kinase, useful for	
XX	identifying modulators, potentially useful as herbicides and growth	
PT	regulators	
PT		

XX Claim 14; Page 13-14; 18pp; German.
 PS
 CC The invention relates to a nucleic acid (1, ABL60244) that encodes a
 CC plant phosphomemalonate kinase (PMVK, ABL77505), excluding the known
 CC fully defined partial sequences ABL60245-ABL60247. Plant PMVK (1),
 CC constructs and host cells that contain (1) are used to identify agents
 CC that bind to and/or modulate activity of PMVK, potentially useful as
 CC herbicides and growth regulators. (1) is also used for recombinant
 CC production of PMVK.
 CC
 XX Sequence 505 AA;
 SO
 Query Match 100.0%; Score 2581; DB 23; Length 505;
 Best Local Similarity 100.0%; Pred. No. 1,7e-236;
 Matches 505; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 0;
 QY 1 MAVVASAPGKYLMTGTYLVEKPNAGLVLSTNAPFAIVKPINEVVPESAMKWTQKL 60
 DB 1 MAVVASAPGKYLMTGTYLVEKPNAGLVLSTNAPFAIVKPINEVVPESAMKWTQKL 60
 QY 61 TSPOLSRSMYKLSLNLTLQSVASDSRNPFEHAIOYATAAHLATEKDESKHL 120
 DB 61 TSPOLSRSMYKLSLNLTLQSVASDSRNPFEHAIOYATAAHLATEKDESKHL 120
 QY 121 QGLDITILGSNDYSTRNQIESAGLPLTPESLGTLPASTITFNAESNGANSKPEVAKT 180
 DB 121 QGLDITILGSNDYSTRNQIESAGLPLTPESLGTLPASTITFNAESNGANSKPEVAKT 180
 QY 181 GLSSAAMTAVVAALLHYLGVDLSDPCKEGKFCGSDLDVTHMAQSHCLAGKXVSSG 240
 DB 181 GLSSAAMTAVVAALLHYLGVDLSDPCKEGKFCGSDLDVTHMAQSHCLAGKXVSSG 240
 QY 241 FDVSCAVYSGSORYRFSPEVLSPFAQVAVTGLPLNEVIGTILKWKDNKRTFESLPLMLNL 300
 DB 241 FDVSCAVYSGSORYRFSPEVLSPFAQVAVTGLPLNEVIGTILKWKDNKRTFESLPLMLNL 300
 QY 301 FLGEPGSGSSTPSMGAVKWKQMSDPEKARENNONLS DANLELETKLNDLSKLAKDHW 360
 DB 301 FLGEPGSGSSTPSMGAVKWKQMSDPEKARENNONLS DANLELETKLNDLSKLAKDHW 360
 QY 361 VYLAVIKSCSVLTSEKWLHATEPINEAIIKELLEAREAMLRIRILMOMGEAASVPIEP 420
 DB 361 VYLAVIKSCSVLTSEKWLHATEPINEAIIKELLEAREAMLRIRILMOMGEAASVPIEP 420
 QY 421 ESQOTLDSTMSAGVLAGVPAGGDFALPALTGLDSTGKLTQAMSSHNVLALLVREDP 480
 DB 421 ESQOTLDSTMSAGVLAGVPAGGDFALPALTGLDSTGKLTQAMSSHNVLALLVREDP 480
 QY 481 HGVCLSESDPRTTCITTCVSSIHLE 505
 DB 481 HGVCLSESDPRTTCITTCVSSIHLE 505
 RESULT 2
 AAU15094
 ID AAU15094 standard; Protein; 432 AA.
 AC AAU15094;
 XX
 DT 04-DEC-2001 (first entry)
 XX
 DE Protein encoded by C. albicans essential gene CayMR220W (ERG8).
 XX
 KM Gene identification: essential gene; GRACE: pathogenic fungus;
 KM gene replacement and conditional expression; fungal infection.
 XX
 OS Candida albicans.
 XX
 PN WO200160975-A2.
 XX
 PD 23-AUG-2001.
 XX

PF 20-FEB-2001; 2001WO-US05551.
 XX
 PR 18-FEB-2000; 2000US-0183534.
 XX
 PA (ELIT-) ELITRA PHARM INC.
 XX
 PI Roemer T, Jiang B, Boone C, Bussey H;
 XX
 DR WPI; 2001-489080/53.
 DR N-PSDB; AAS23422.
 XX
 PT Identifying genes essential to fungal metabolisms and identifying
 PT potential therapeutic agents that target these genes -
 PS Claim 43; Page 226-227; 324pp; English.
 XX
 CC The present invention relates to novel methods for constructing fungal
 CC strains useful for identification and validation of gene products as
 CC targets for therapeutic agents, for creating a collection of identified
 CC essential genes, and screening assays for the discovery of new drugs.
 CC The invention provides the GRACE (gene replacement and conditional
 CC expression) method for the construction of mutant organisms referred to
 CC as GRACE strains of the organism. The invention can be applied to any
 CC organism, particularly a pathogenic fungus e.g. Candida albicans,
 CC Aspergillus fumigatus and Cryptococcus neoformans. The methods are
 CC useful to identify agents that may be used in the treatment of fungal
 CC infections. AAU15053-AAU15113 represent proteins encoded by C. albicans
 CC essential genes.
 XX

SO Sequence 432 AA;
 Query Match 18.6%; Score 480; DB 22; Length 432;
 Best Local Similarity 28.5%; Pred. No. 1.4e-36;
 Matches 144; Conservative 78; Mismatches 183; Indels 100; Gaps 15;
 0;
 QY 1 MAVVASAPGKYLMTGTYLVEKPNAGLVLSTNAPFAIVKPINEVVPESAMKWTQKL 60
 DB 1 MSAKAPSPAGKFLAGGLVLEPIYDAYVTALSSMHAHTPKISLSE-----SRKI 53
 QY 61 TSPOLSR-ESMYKLSLNLTLQSVASDSRNPFEHAIOYATAAHLATEKDESKHL 119
 DB 54 SSPFANGEMETHLSSN--TEKPEVOSRINPFEATIFIVLAVIOPTLEARD----- 103
 QY 120 QGLDITILGSNDYSTRNQIESAGLPLTPESLGTLPASTITFNAESNG-----AN 172
 DB 104 ---LEIIYSDPGYHQED-----TETKTSNMGKTFLYHSR 137
 QY 173 SKPEVARTGSSAAMTAVVAALL-HYL-GVVDLSDPCKEGKFCGSDLDVTHMAQSH 230
 DB 138 ALTEVEKTLGSSAGLVSAVTSLSHFIPNVIS-----TNKDLHVAQIAH 185
 QY 231 CLAOGKVGSGFDVSCAVYSGSORYRFSPEVLSPFAQVAVTGLPLNEVIGTILK---GKMD 286
 DB 186 CYACKTKIGSGFDVATATYGSIVYRRFOPALINDVQYLESDP--EKPELTAKLIASWE 243
 QY 287 NKRTFESLPLMLNLFLGEPGSGSSTPSMGAVKWKQMSDPEKARENNONLS DANLELE 346
 DB 244 FKHERCTLPNGIKLIMGDV-KGGSETPKLYSRVLOMKKEESSVYVDOLNSANLFMK 302
 QY 347 KLNDLSKLAKDHWVYLVKIKSCSVLTSEKWLHATEPINEAIIKELLEAREAMLRIRIL 406
 DB 303 ELREMKREKYSDPETIYKED-----HSEPLVYAL-----KNIRKG 339
 QY 407 MROMGEAASVPIEPESOTLDSTMSAGVLAGVPAGGDFALPALTGLDSTGKLTQAM 466
 DB 340 LQALTQKSEVIEPDVOTGLDRQCEIPGCGVVPAGGVDALAVLLEHQVNFQOKT 399
 QY 467 SS-----HNVALLVREDPHGVCLE 486
 DB 400 LENDPYFHNMYWVDLEBQTEGVLEE 424
 RESULT 3

AA72679
 ID AA72679 standard; Protein: 432 AA.
 AC AA72679;
 DT 31-MAY-2001 (first entry)
 XX
 DE Candida albicans phosphomevalonate kinase (PMK; ERG8) protein.
 XX
 KM Phosphomevalonate kinase; PMK; ERG8; anti-fungal agent; diagnosis;
 KM Infection.
 OS Candida albicans.
 XX
 FH Key location/Qualifiers
 FT Misc-difference 244
 FT /note- "Encoded by TTC of the sequences shown in
 FT AAD02791 and AAD02792"
 XX
 PN MO200114533-A2.
 XX
 PD 01-MAR-2001.
 XX
 PF 15-AUG-2000; 2000WO-GB03100.
 XX
 PR 21-AUG-1999; 99GB-0019766.
 XX
 PA (ASTR) ASTRAZENECA AB.
 PA (ASTR) ASTRAZENECA UK LTD.
 PI Rosamond JDC, Schnell NF;
 XX
 DR WPI: 2001-218441/22.
 DR N-PSDB: AAD02791, AAD02792.
 XX
 PT New polypeptides and polynucleotides (ERG8) from Candida albicans,
 PT useful in assays for identifying inhibitors of phosphomevalonate kinase
 PT activity and as reagents for diagnosing C. albicans infection -
 XX
 PS Claim 1; Page 27-28; 29pp; English.
 XX
 CC The present sequence is phosphomevalonate kinase (PMK; ERG8) protein
 CC from Candida albicans. The ERG8 protein is useful in an assay for
 CC identifying compounds that inhibit phosphomevalonate kinase (PMK)
 CC activity. These inhibitors are useful as anti-fungal agents. The ERG8
 CC DNA and protein are also useful as reagents for diagnosing C. albicans
 CC infection.
 CC
 XX
 SO Sequence 432 AA;
 Query Match 18.4%; Score 475; DB 22; Length 432;
 Best Local Similarity 28.3%; Pred. No. 4.3e-36;
 Matches 143; Conservative 78; Mismatches 184; Indels 100; Gaps 15;
 QY 1 MAVVASAGKYLMTGVLLEKPNAGLYLSTNARFYATKPNIEKPKESMAKWTVDYL 60
 DB 1 MSKAFSAGKAFLAGGYLVLEPIYDAYVTALSSRMHAYITPGTSLKE-----SRKI 53
 QY 61 TSPQISR-ESMYKLSLNHLTQSVASDSRNPFEHAIOYALAAHNLATEKKESLKIL 119
 DB 54 SSPQANEMEWETHISSN--TEKPREVQSRINPLEATITFIYAIQPTFAFD----- 103
 QY 120 LQGLDITLIGSNDFSYRNQIESAGLPLTPESLGLAPFASITFMAESNG-----AN 172
 DB 104 ---LEITIIYSDPGYHSQED-----TEKRTSSNGEKTELYHSR 137
 QY 173 SKPEVAKTGLSSAAMTTAAVVAALL-HYL-GVVDLSDPCKEKGFCSDLDVTHMIAGTSH 230
 DB 138 ATTEVEKTGLSSAGLVASVATSLSHFLPNVIS-----TNKDILHNAQIAH 185
 QY 231 CLAQKVGSGPVCAYVGSQRYVFESEVLSFAOVAAVTGLPLNVEYITLTK----GKWD 286
 DB 186 CYAOKKISSGPDVAIVGLYVRRFOPALINDVQVLESDE--EKFPTEKLKLIESNWE 243

QY 287 NKRTPESLPPLNLFLGEPGSGSSTPSWGAVKKKWMSDPEKARENMONLSANLELET 346
 DB 244 EKHERCTLPYGIKLMGDV-KGSETPKLVSRVLOWKKEKPESSVYIDQASAWLQPMK 302
 QY 347 KLNDLSKLAKDMVDVLYRYIKSCSVLTSEKWLHATEPINEAIKELLEAREAMLRIRIL 406
 DB 303 ELREKREKYDSDPEYIKELD-----HSVEPLTVAI-----KNIRKG 339
 QY 407 MRQGEAASVPIEPESQOLDSTMSAGCVLLAGVPGAGGFPAITPAITLGDSTLTQAM 466
 DB 340 LQALTOKSEVPTEPDVQOTLDRCOEIPGCVGVPAGAGYDAIVLVLENQVGNFKQRT 399
 QY 467 SS-----HNVLLALVREDPHCYCLE 486
 DB 400 LENPDYFHNVYWDLEBOTEGVLEE 424
 RESULT 4
 ID AAB18130 standard; Protein: 109 AA.
 XX
 AC AAB18130;
 XX
 DT 08-NOV-2000 (first entry)
 XX
 DE Pinus radiata phosphomevalonate kinase protein SEQ ID NO:291.
 XX
 KM Eucalyptus grandis; Pinus radiata; modification; isoprenoid; plant;
 KM metabolism; isoprenoid biosynthetic pathway; terpenoid; steroid;
 KM genome mapping; physical mapping; positional cloning; forestry;
 KM agriculture; medicine; fermentation; plant development; pest resistance;
 KM pinene; myrcene; Monterey pine.
 XX
 OS Pinus radiata.
 XX
 PN WO200036081-A2.
 XX
 PD 22-JUN-2000.
 XX
 PF 16-DEC-1999; 99WO-NZ00219.
 XX
 PR 17-DEC-1998; 98US-0215504.
 PR 29-JUL-1999; 99US-0146441.
 XX
 PA (GENE-) GENESTIS RES & DEV CORP LTD.
 PA (FLET-) FLETCHER CHALLENGE FORESTS LTD.
 XX
 PI Havukkala IJ;
 XX
 DR WPI: 2000-431575/37.
 DR N-PSDB: AAA69566.
 XX
 PT New plant polynucleotides encoding polypeptides involved in the
 PT production and modification of isoprenoids, useful in forestry and
 PT agriculture for manipulation of isoprenoid metabolism -
 XX
 PS Claim 26; Page 159; 164pp; English.
 XX
 CC The present invention describes plant polynucleotides encoding
 CC polypeptides involved in the production and modification of isoprenoids,
 CC such as terpenoid and steroid compounds. The polynucleotides are used
 CC in genome mapping, in physical mapping and in positional cloning of
 CC genes. The polynucleotides and polypeptides are useful in forestry and
 CC agriculture for manipulation of isoprenoid metabolism, in medicine for
 CC therapeutic effects, including direct application in diseased organisms
 CC or indirect application by transgenic organisms and in fermentation and
 CC chemical processing industries involving isoprenoids. In plant
 CC applications, manipulating isoprenoid pathways or isoprenoid composition
 CC may, for example, affect plant development, pest resistance, and the
 CC value of extractives (e.g. pinene and myrcene). The ubiquitous and
 CC varied roles of isoprenoids make the polynucleotides attractive targets
 CC for biotechnical applications in a variety of fields. AAA69527 to

CC AAA69690 and AAB18004 to AAB18143 represent *Eucalyptus grandis* and *Pinus*
CC *radiata* polynucleotides and proteins used in the exemplification of the
CC present invention.
XX
SQ Sequence 109 AA;

Query Match 14.4%; Score 372.5; DB 21; Length 109;
Best Local Similarity 68.6%; Pred. No. 2.5e-27;
Matches 72; Conservative 15; Mismatches 17; Indels 1; Gaps 1;

QY 1 MAVVASAPGKVTMTGTYLLEKPNAGLVLTSTNAREFAIVKRIINEEVPESMAWKTDVKL 60
Db 1 MAVVVSAPGKLTITGALLILEKPNPGLVLTARFAIVKRLKRTSDSSNAMLMTDVKL 60

QY 61 TSPOLRESMYKLSLNHLTLOSASDSR-NPEVEHAIOYAIAAA 104
Db 61 TSPOLAKEAIVKLSLKTLISQNVASSSSNGNPFEQAVQFVAAA 105

RESULT 5
AAU35063
ID AAU35063 standard; protein; 368 AA.

AC AAU35063;

DT 13-FEB-2002 (first entry)

DE Enterococcus faecalis cellular proliferation protein #350.

KM Antisense; prokaryotic cellular proliferation protein;

KW antibiotic; antibacterial; drug design.

OS Enterococcus faecalis.

PN WO200170955-A2.

PD 27-SEP-2001.

PE 21-MAR-2001; 2001WO-US09180.

PR 21-MAR-2000; 2000US-191078P.

PR 23-MAY-2000; 2000US-206848P.

PR 26-MAY-2000; 2000US-207272P.

PR 23-OCT-2000; 2000US-242578P.

PR 27-NOV-2000; 2000US-253625P.

PR 22-DEC-2000; 2000US-257931P.

PR 16-FEB-2001; 2001US-269308P.

XX (ELIT-) ELITRA PHARM INC.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

CC for homologous nucleic acids which are required for cell proliferation in
CC a wide variety of organisms. The present sequence represents an
CC essential prokaryotic cellular proliferation protein.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 368 AA;

Query Match 6.9%; Score 178.5; DB 22; Length 368;
Best Local Similarity 22.3%; Pred. No. 5.6e-08;
Matches 106; Conservative 58; Mismatches 179; Indels 133; Gaps 20;

QY 3 VVASAPGKVTMTGTYLLEKPNAGLVLTSTNAREFAIVKRIINEEVPESMAWKTDVKL 62
Db 2 IEVTPGKLFAGHAYVEPCHPAIIVADQFVVEETFEEDSIOASQSYSPIRMT- 60

QY 63 POLRESMYKLSLNHLTLOSASDSR-NPEVEHAIOYAIAAAHLATEKDESLKLLLOG 122
Db 61 --RRNGELVLDLR-----ENPF-----HYVLAAILH-TEKVAQEQNKEL--- 96

QY 123 LDTITLGSNDPFSYRNQIESAGLPLPESLGLTAPFASITFENNAESNGANSKPEVAKTGL 182
Db 97 -----SEYHLK-----VTSELDSSNG-----RKYGL 117

QY 183 GSSAAMTTAAVAAL-LHYLGVDLSDPCKEKGFGCSDLDTVIMIAOTSCLAGKVGSG 240
Db 118 GSSGAVTVGTYKALINIFYDGL-----ENEIFPLSALAHVQGN-GSC 161

QY 241 FDVSCAVYGSORVYREPEVLSEFAQNAVGTGLPLNEVIGTILKGMQNKRT-ESLPLPM 299
Db 162 GDIVASCYGG--WIAFTSFHDWVNOKVT---TETLTDLADWDPELMTEPLKVPQRLR 215

QY 300 LFLGEPGSGSPTPSMAYGAVKKMOMDPEKAREKQMLSDANLELETKLNDLSKLADHW 359
Db 216 LLIIGTGS-FASTSDLVDRV---HQSKKEQAAEQGLMSRLCVERMINGFN----- 264

QY 360 DYLIRVAKSCSVLTSEKWLHATEPINEAIIKELLEAREMLRIRILRMQGEAASVPTE 419
Db 265 -----TGKIS-VIOKQITKNRQ-----LAEISLITGVIE 294

QY 420 PESOTQLDSTMSAEGVLAGVPGAGGEDAIFATLGDG-TRLTQWSSHNVLAL 474
Db 295 TEALKNLCDLAEISTGA--AKSSGAGGDCGIYIFROKSGILPLMTAMERDGIPL 348

RESULT 6
AAB60859
ID AAB60859 standard; protein; 368 AA.

AC AAB60859;

DT 30-MAR-2001 (first entry)

DE Mevalonate pathway protein #3.

KM Mevalonate pathway; disease; infection.

KW Enterococcus faecalis.

OS Enterococcus faecalis.

PN WO200078935-A1.

PD 28-DEC-2000.

PR 22-JUN-2000; 2000WO-US17262.

PR 22-JUN-1999; 99US-0140519.

PR 02-AUG-1999; 99US-0146682.

XX (SMIK) SMITHKLINE BEECHAM CORP.

XX (SMIK) SMITHKLINE BEECHAM PLC.

XX

XX

XX

XX

XX

XX

XX

XX

PI Brown JR, Gwynn M, Mathie TB, Myers JE, Trainl CM, Van Horn S;
PI Wilding EJ;
XX
DR WPI: 2001-071392/08.
XX
PT New isolated mevalonate pathway gene polynucleotide derived from
PT bacterium is useful for treatment of bacterial infection -
XX
PS Claim 20; Page 14; 158pp; English.
XX
CC The present invention relates to an isolated mevalonate pathway gene
CC derived from a bacterium from clade of Class II of the
CC phylogenetic tree referred to in the specification. The invention may be
CC used for treatment of disease related to bacterial infection, e.g.
CC conjunctivitis, pneumonia, bacteremia and meningitis.
XX
SQ Sequence 368 AA;

Query Match 6.9%; Score 178.5; DB 22; Length 368;
Best Local Similarity 22.3%; Pred. No. 5.6e-08;
Matches 106; Conservative 58; Mismatches 179; Indels 133; Gaps 20;

QY 3 VVAAPGKVLMTGGLVLEKPNAGLVLSTNARFAIVPINEEKPESWAMKMDVKLTLS 62
DB 2 IEVTPPGKLTAGEAVAVPEGPALIVAVDQFVTYVEETIDEGSIQAOYSLSPIRMT 60
QY 63 POLRESMYKLSLNHLTLQSVASDSRNPFEHAIOYIAAHLATEKDKSLHKLLQ 122
DB 61 ---RRNGELVDIR-----ENPF-----HYLAAILH-TEKYAQBNKEL--- 96
QY 123 LDITLLSNDFYSYRNOIESAGLPLTPESLGTLPFASITFNMAESNGANSKPEVAKTGL 182
DB 97 -----SEYHLK-----VTSELDSSNG-----RKYGL 117

QY 183 GSSAAMTAVVAAL--LHYLGVDLSDPCKEGKFCGSDLDVTHMAQSHCLAQKVGSG 240
DB 118 GSSGAVTVGTAKALNIEFDLGL-----ENEFIFKLSALHLAVQGN-GSC 161
QY 241 FDVCAVYGSORVYRFSPEVLSPFAOVAVTGLPLNEVIGITLKGKDNKRT-EFSPLPLMN 299
DB 162 GDIAASCGG--WIAFSFHDHWNOKYT-----TETLDDLAMPPELMIFPKYPKDLR 215
QY 300 LFLGPGSGGSTPSMGAVKRWQMSDEKAREMNQNSDANLETRKLDLSKLAKDHW 359
DB 216 LLIGWTGS-PASTSDLVDRV---HOSKEKQAAVEQFLMKSLRCVETMINGFN----- 264
QY 360 DYLRLVTKSCSVLSEKVNLAITEPINEAIIKELIAREAMLRIRILMKRQGEAASVPIE 419
DB 265 -----TGKIS-VIOKQITRNO-----LAEELSLTGVAIE 294

QY 420 PESQTQLDSTMSAEGVLLAGVPGAGFDALFAITLGDGSG-TKLTOAMSSHNVAL 474
DB 295 TEALKNLCDLAEYSTGA--AKSSGAGGDCGIVIFROKSGILPLMTAMEKQGITPL 348

RESULT 7
AAB99728
ID AAB99728 standard; Protein: 374 AA.
XX
AC AAB99728;
XX
DT 10-SEP-2001 (first entry)
XX
DE Streptomycetes sp. CL190 mevalonate pathway orfC protein SEQ ID NO:10.
XX
KW Streptomycetes sp. CL190; mevalonate pathway; actinomycetes; cardiant;
KW isoprenoid compound; osteopathic; cytoskeletal; ubiquitons; vitamin K2;
KW carotenoid; heart disease; osteoporosis; cancer; drug; health food.
XX
OS Streptomycetes sp. CL190.
XX
PN MO200142476-A1.

PD 14-JUN-2001.
XX
PF 06-DEC-2000; 2000WO-JP08620.
XX
XX 08-DEC-1999; 99JP-0348375.
XX
PA (SETO/) SETO H.
PA (KUZU/) KUZUYAMA T.
XX
PI Seto H, Kuzuyama T, Takahashi S, Takagi M;
XX
DR WPI: 2001-381696/40.
XX
PS N-PSDB: AAH44046.
XX
PT Actinomycetes-originated genes of enzymes participating in mevalonate
PT pathway, applicable in producing e.g. ubiquitons, vitamin K2 and
PT carotenoids for treatment of heart diseases, osteoporosis and cancer in
PT drug and health food -
XX
PS Claim 6; Page 62-64; 75pp; Japanese.
XX
CC The sequence given in AAH44043 represents a DNA sequence isolated from
CC Streptomycetes sp. CL190, containing a 6798 base pairs (51), which encodes
CC the whole enzyme necessary for functioning the mevalonate pathway. The
CC sequence encodes protein sequences, designated orfa to E and hmgi, which
CC are used in the mevalonate pathway. The proteins and polynucleotide
CC sequences encoding them have cardiant, osteopathic and cytoskeletal
CC activities. The genes are applicable in producing e.g. ubiquitons,
CC vitamin K2 and carotenoids which can be used in the treatment of heart
CC diseases, osteoporosis and cancer in drugs and health foods. The present
CC sequence represents the orfC protein from the present invention.
XX
SQ Sequence 374 AA;

Query Match 6.0%; Score 155; DB 22; Length 374;
Best Local Similarity 19.7%; Pred. No. 1e-05;
Matches 97; Conservative 68; Mismatches 172; Indels 156; Gaps 19;

QY 3 VVAAPGKVLMTGGLVLEKPNAGLVLSTNARFAIVPINEEKPESW 51
DB 8 IYRANPGKLTAGEAVAVDQFPGPALIVAVDRHISTVSDADADICADAVYISSDLGPAY 67
QY 52 AMKWDVRLTSPOLRESMYKLSLNHLTLQSVASDSRNPFEHAIOYIAAHLATEKD 111
DB 68 GWRWHDGLV-----VNDP--DDQQAASLAHV----- 95
QY 112 KESLHKLLQGLDITLLSNDFYSYRNOIESAGLPLTPESLGTLPFASITFNMAESNGA 171
DB 96 -----SAIEYVGRLLGGRGOKVPALTLTSSRLHEDG-- 127
QY 172 NSKPEVAKTGLSSAAMTAVVAALHLYGV-VYDLSDECKEKGKFCGSDLDVTHMAQSH 230
DB 128 -----RKFGLGSSGAVVAVVAALAAFCGLSTDEFRILAMLATAYELD----- 171
QY 231 CLAQKVGSGFDVSCAVYGSORVYRFSPEVLSPFAOVAVTGLPLNEVIGITLKGKDNKRT 290
DB 172 -----PKSGGDLAASWTGG--WIAVQAPDRAF-----YLDLARRGVDRTLKAPPGHSV 220
QY 291 EFSLPLMLNPL-----GEPGSGGSTPSMGAVKRWQMSDEKAREMNQNSDANLELE 345
DB 221 R-RLPAPKGLTLEVGWTEGP-----ASTASLVS-----DLHRRTRWGSASQREVE 265
QY 346 TKLNDLSKLARDHMDVYLRVTKSCSVLSEKVNLAITEPINEAIIKELIAREAMLRIRI 405
DB 266 T-----TTDCVRSAYTALESGB--PTSLHETRRARQELAR--- 299
QY 406 IMRQNGEAASVPIEPESQTQLDSTMSAEGVLLAGV-PGAGFDALFAITLGDGSGTKLT- 463
DB 300 -----LDDEVGIGITFPKLTALCD--AAEAVGAAKPSGAGGDCGIALDLDAEASRDLTH 352
QY 464 --QAMSSHNVAL 474
DB 353 VRORWETAGVLP 365

CC conjunctivitis, pneumonia, bacteremia and meningitis.
XX
SQ Sequence 361 AA;

SQ Sequence 361 AA;

Query Match	5.7%;	Score 147.5;	DB 22;	Length 361;
Best Local Similarity	21.2%;	Pred. No. 4.9e-05;		
Matches 102;	Conservative 64;	Mismatches 171;	Indels 145;	Gaps 19

QY	3	VVAAPAGCVLMTGGVILVLEKRNALCVSTNARFAIYKPTINEEVKPRPSMAMKMTDVLTLS	62
Db	2	IEVSNAPKGLIYAGEYAVVEIGHRAVIAANDFOYVTVYESARKKGSIOSAOTSGMPAWT-	60
QY	63	POLRESMWKLSLNLHTLQSVASDSRNPFVEHAIOYAAIAHLATEKDESLHKLLOG	1222
Db	61	---RRNGELVDIR-----ENPF-----HYLAAIRL-TEKYAQEKNIL---	96
QY	123	LDITILGSDNPFYSTRNOIESAGLPLTFESLGTLPAPFASITFMAESGANSKEPVAKTGL	1622
Db	97	-----SEYDLK-----VTSELDSSNG-----RKYGL	1177
QY	183	GSSAAMTAAVVAALLHYLVGYVDLSDPCKEKGFGSGSDIDVTHIMTQSHCLAQGVSGGFD	2422
Db	118	GSSGAVYATYAKALNVFAL-----NLQIEFKRIALLAN--LAVDNGSGCGD	1633
QY	243	VSCAVYGSQRYVRESPEVLSFAQYAVTGLPLNEVITGTLGKMDNKRTEESLPPL-----	2977
Db	164	IAASCYGS- -WIASTFDPHPLQDEQPHSHSELALDWG-----LSIEPLINPED	2133
QY	298	MNPLFLGEGSGSGSSTPSMVGAVKWKQMSDPEKAEENNONISDANLELETKLNDLSKLAKD	3577
Db	214	LRLDIGWGS- -PASTSDLVDPV-----HRSRED-----	2400
QY	358	HMDVYLVRKISCVLTSEKWLHLATEPINEAI--IKE--LLEAREAMLRIRILRMONGEA	4133
Db	241	-----KNVATQFLKNSTECVNEIKFKFENNVTLLIQMTRKNROLHLHLSAI	2888
QY	414	ASVPIEPESQTOLDSTMSAGCVLLAGVPAGAGFDAIFATITLGDG--TKLTQAMSSSHVL	4722
Db	289	TGVVIEFPLAKKLCLNLAEQYEGA--AKSSGAGGGDCGIVIDOKSGILPLMSAMEKAEIT	3466
QY	473	AL 474	
Db	347	PL 348	
RESULT 10			
ID	AMG81912	standard; Protein: 358 AA.	
XX	AMG81912;		
XX	03-SEP-2001	(first entry)	
DE	S. epidermidis open reading frame protein sequence SEQ ID NO:918.		
KW	Staphylococcus epidermidis srt1 strain; infection; diagnosis;		
KW	vaccination; endocarditis.		
OS	Staphylococcus epidermidis.		
PN	MO200134809-A2.		
PD	17-MAY-2001.		
PR	09-NOV-2000; 2000MO-0330782.		
PR	09-NOV-1999; 99US-0164258.		
PA	(GLAX) GLAXO GROUP LTD.		
PI	Kimmerly MJ;		
OR	WPI; 2001-316495/33.		

DR N-PSDB; AAH52762.

PT Nucleic acids encoding polypeptides from *Staphylococcus epidermidis*
PT useful for vaccinating against infections, e.g. endocarditis -
XX
PS Claim 18; Page 273; 2186pp; English.

AAH5230 to AAH53970 represent nucleic acids (I) encoding polypeptides (II), given in AAG81454 to AAG83120, from *Staphylococcus epidermidis*. (I) and (II) can have antibacterial activity and therefore can be used in vaccination. The nucleic acids (I) may be used to produce the S. epidermidis polypeptides (II) via the production of vectors containing them which are used to produce host cells which express the polypeptides. The polypeptides (II) (and/or nucleic acids) may then be used to vaccinate subjects and to raise antibodies against the bacteria. The polypeptides may also be used to assay for other inhibitors of their activity and therefore identify compounds that may be used for the treatment of S. epidermidis infections, e.g. endocarditis. AAH53971 to AAH55090 represent specifically claimed S. epidermidis genomic DNA polynucleotide sequences from the present invention. AAH55091 to AAH55098 represent oligonucleotide sequences and primers which are used in the exemplification of the present invention.

N.B. The present invention specifically claims all the polynucleotide sequences given in the sequence listing of the present specification, however the sequence listing only goes up to SEQ ID NO:4454 so even though sequences are given in the disclosure for SEQ ID NO:4465 to 4472 no sequences are present for SEQ ID NO:4455 to 4464.

SQ Sequence 358 AA;

Query Match	5.5%;	Score 142.5;	DB 22;	Length 358;
Best Local Similarity	19.8%;	Pred. No. 0.00014;		
Matches 94;	Conservative 72;	Mismatches 167;	Indels 141;	Gaps 20

Oy	3	VVASPGVIMTGGVLYLEKRNAGIVYSTNRPAFAIVPINEEVKPESMAMKMTDVKLT	62
Oy	2	IQVKKPGLYLTAGEVAVTEPGYKSTLLIANPRTTITASK-----VEGSIS	50
Oy	63	POLARESM-YKLSLNHLTLOSASDSRNPVEHAIOYAIAAHLATEKDESHKLILQ	122
Db	51	KLHAEYKPFDRNEDRIETSDVOAKQ-----LKVVAIAIEV-----	87
Oy	122	GLDITILGSNDPFYSYRNOIESAGLPPLRESGLTAPLPAFASITFNAESGANSKEPVAKTG	18
Db	88	-----FEQYVRSNNMLKHFHL-----TIDSNLADNSG-----QKYG	11
Oy	182	LGSNAAMTTAVYAALHLHLYGVYDLSDPCKEKGFGCSLDIVIHMLAOTSCHLAOGKVGSGF	24
Db	120	LSSSAAVAVSVKALANEYFG--LELSNL-----YLYKLAIVANMKLOSLSGCG--	16
Oy	242	DVSCAVY--GSDRYVFESEPVLSFQOVAVTGLPLNEVIGTILKGMKDKRTE--FSLPLMN	299
Db	166	DIANSVYSGMLATYFEDHDWK--QOMETS-----VADVLEKMPGLHTEPLQAPENME	21
Oy	300	LFLGPGSGGSSPTSMGAVKRWQMSDEKARENMQNLSDANLEETKLANDSKLANDHW	355
Db	219	VLIGMTGSPASS--PHLYSEVRLK--SDP-----SFY	24
Oy	360	DVYLRLVYKSCSVLISEKKVWLHATEPINEAIIKELLEAREAMLR--RIIMQMGPAASVPI	41
Db	248	GDFLDOSSHAC-----VESLLOAFPTNNIKGVQKMIKIRNRKIRLOSMDNEASVEI	295
Oy	419	EPESOTOLDSTMSAEGVLLAGVPGAGGFDAIPAITLGDSGTKLQOAMSSHNVL	472
Db	296	ETDKKTKLCVDYGEKHGGA--SKTSGAGG-----GDCGITIINKVYIDKNIT	338

Db 170 DIAVSYSGLMAYSTFDHDKWK-QOMETS-----VNDVLEKMPGLHIEPLQAPENME 222
QY 300 LFLGPPGGSGSTPSMGAVKKQMSDEPKARENMQNSDANLELETNLSKLAKDHW 359
Db 223 VLIQGTGSPASS-PHLVSEVRRLK-SDP-----SFY 251
QY 360 DVLRLVIRKSCSVLTSEKMWLHATEPINEALIKELLEAREAMLRI-RILROMGEAASVPI 418
Db 252 GDFLDQSHAC-----VESLLOAFKTNKYGQKMRIRKRITIOSMDEASVEI 299
QY 419 EPESOTQLDSTMSAEGVLLAGVPGGPDALFAITLGDSCGTLQAMSSHNVL 472
Db 300 ETDKRLKLDYGEKHGA--SKTSGAGS-----GDCGTTITNKVIDKNII 342

RESULT 13

AAB60864 standard; protein: 358 AA.

AAB60864;

30-MAR-2001 (first entry)

Mevalonate pathway protein #2.

Mevalonate pathway; disease; infection.

Staphylococcus aureus.

W0200078935-A1.

28-DEC-2000.

22-JUN-2000; 2000WO-US17262.

22-JUN-1999; 99US-0140519.

02-AUG-1999; 99US-0146682.

(SMIK) SMITHKLINE BEECHAM CORP.

(SMIK) SMITHKLINE BEECHAM PLC.

Brown JR, Gwynn M, Mathie TB, Myers JE, Trainl CM, Van Horn S;

WPI; 2001-071392/08.

New isolated mevalonate pathway gene polynucleotide derived from

bacterium is useful for treatment of bacterial infection -

Claim 20; Page 17-18; 158pp; English.

The present invention relates to an isolated mevalonate pathway gene

derived from a bacterium from clade of Class II of the

phylogenetic tree referred to in the specification. The invention may be

used for treatment of disease related to bacterial infection, e.g.

conjunctivitis, pneumonia, bacteremia and meningitis.

Sequence 358 AA;

Query Match 5.3%; Score 136.5; DB 22; Length 358;

Best Local Similarity 21.5%; Pred. No. 0.00054;

Matches 103; Conservative 60; Mismatches 180; Indels 135; Gaps 23;

QY 120 LOGIDITILGSDNFYSYRNOIESAGLPLTPESLGLTAPASTTFMAESNGANSKPEVAK 179

Db 91 -----YAKSCDIAMKHFHLTIDS-----NLDSNG-----HK 117
QY 180 TGLGSSAMTTAVVAALLHYLVNLSDPCKEKGCCDDVYIHIAOTSICLAGKXGS 239
Db 118 YGLGSSAVALVSVKLVNEFDPM-----KLSNL-YIYLAIVANNKLOSLSGC 164
QY 240 GFDVCAVY-GSORVRRSPEVLSPAOYAVTGLPLNEVIGTILKQKMNKRTF-FSLPPL 297
Db 165 G-DIAVSYSGLMAYSTFDHDKWK-QOMETS-----VNDVLEKMPGLHIEPLQAPEN 216
QY 238 MNLFLGPPGGSGSTPSMGAVKKQMSDEPKARENMQNSDANLELETNLSKLAKDH 357
Db 217 MEVLIGMTGSPASS-PHLVSEVRRLK-SDP-----SFY 251
QY 358 HMDVLRVIRKSCSVLTSEKMWLHATEPINEALIKELLEAREAMLRI-RILROMGEAASVPI 418
Db 246 FYGDFLEDSSHRC---VLR-LIHAFTNNIKGVQKMRVQN-----RPIQRMDEATVD 294
QY 418 EPESOTQLDSTMSAEGVLLAGVPGGPDALFAITLGDSCGTLQAMSSHNVL 472
Db 295 IETKRLKLDYGEKHGA--SKTSGAGGDCGTTITNKVIDKNII 342

RESULT 14

AAU37273 standard; protein: 358 AA.

AAU37273;

14-FEB-2002 (first entry)

Staphylococcus aureus cellular proliferation protein #1443.

Antisense; prokaryotic cellular proliferation protein;

antibiotic; antibacterial; drug design.

Staphylococcus aureus.

W0200170955-A2.

27-SEP-2001.

21-MAR-2001; 2001WO-US09180.

21-MAR-2000; 2000US-191078P.

23-MAY-2000; 2000US-206848P.

26-MAY-2000; 2000US-207727P.

23-OCT-2000; 2000US-242578P.

27-NOV-2000; 2000US-253625P.

22-DEC-2000; 2000US-257931P.

16-FEB-2001; 2001US-263088P.

(ELITR) ELITRA PHARM INC.

Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;

Yamamoto RT, Xu HH;

WPI; 2001-611495/70.

N-PSDB; AAS55132.

New polynucleotides for the identification and development of

antibiotics, comprise sequences of antisense nucleic acids -

Example 3; Seq ID No 12866; 51pp; English.

The invention relates to antisense inhibitors of genes essential to

prokaryotic cellular proliferation, their use in identifying the

genes, their use in the discovery of novel antibiotics, the essential

genes themselves and the encoded proteins. The prokaryotes used are

Wed May 7 14:13:42 2003

us-09-988-863a-2.rag

Page 11

Db 248 LEDSHRC---VEK-LIHAFTNNIKGVOKKVRON-----R11QRMDEATVDIETEK 296
QY 423 QTOILDSTMSAEGVLLAGVPGAGGEDAIFAITLGD-SGTKLTQAMSSHN 471
Db 297 LKYLCDIAEKYHGA--SKTSGAGGDCGTFITINKVDKEDIYDEWTKHGI 344

Search completed: April 26, 2003, 12:55:05
Job time : 42 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 26, 2003, 12:54:20 ; Search time 17 Seconds

(Without alignments)
874.034 Million cell updates/sec

Title: US-09-988-863A-2

Perfect score: 2581

Sequence: 1 MAVVASAPGKVLMTGGLVLT.....ESGDPPTTCITSGVSIHLE 505

Scoring table:

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: Issued_Patents_AA.*

2: /cgn2_6/ptodata/1/1aa/5A.COMB.pep.*

3: /cgn2_6/ptodata/1/1aa/5B.COMB.pep.*

4: /cgn2_6/ptodata/1/1aa/6A.COMB.pep.*

5: /cgn2_6/ptodata/1/1aa/6B.COMB.pep.*

6: /cgn2_6/ptodata/1/1aa/Backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	142.5	5.5	362	4	US-09-134-001C-5355
2	117	4.5	396	6	US-09-134-001C-5355
3	115	4.5	432	4	US-09-306-595C-8
4	108.5	4.2	1477	1	US-08-038-682-4
5	108.5	4.2	1477	1	US-08-302-832-4
6	108.5	4.2	1477	2	US-08-530-198-4
7	108.5	4.2	1477	2	US-08-469-880-4
8	108.5	4.2	1477	2	US-08-728-470-4
9	108.5	4.2	1477	2	US-08-617-697-4
10	108.5	4.2	1477	2	US-08-617-697-4
11	108.5	4.2	1477	4	US-09-206-942-71
12	104.5	4.0	441	4	US-09-651-200-4
13	104.5	4.0	770	1	US-08-445-135-2
14	101.5	3.9	534	4	US-09-651-200-6
15	101.5	3.9	534	4	US-09-651-200-6
16	101.5	3.9	1024	4	US-09-562-737-49
17	101	3.9	1012	2	US-08-475-891A-4
18	101	3.9	1025	2	US-08-567-375-4
19	101	3.9	1025	2	US-08-587-680A-4
20	97	3.8	863	4	US-08-238-303-11
21	95.5	3.7	1294	4	US-08-819-288-3
22	95.5	3.7	1294	4	US-09-400-348-3
23	95.5	3.7	1321	5	US-08-261-822A-3
24	95.5	3.7	1321	5	PCT-US95-07744A-3
25	95.5	3.7	1599	2	US-08-617-697-9
26	95	3.7	1093	4	US-09-315-793-52
27	93.5	3.6	1536	4	US-09-206-942-67

28	93	3.6	608	4	US-08-843-572E-2	Sequence 2, Appli
29	92.5	3.6	640	4	US-09-177-349-4	Sequence 4, Appli
30	91.5	3.5	1013	4	US-09-415-522-8	Sequence 8, Appli
31	91	3.5	500	6	5486473-2	APPLICANT: FUK
32	91	3.5	730	4	US-09-398-865A-2	Sequence 2, Appli
33	91	3.5	730	4	US-09-710-714-2	Sequence 2, Appli
34	90.5	3.5	587	4	US-09-020-465-2	Sequence 2, Appli
35	90.5	3.5	634	2	US-09-020-465-2	Sequence 2, Appli
36	90.5	3.5	634	3	US-09-197-659-2	Sequence 2, Appli
37	90	3.5	1285	1	US-07-582-945-2	Sequence 2, Appli
38	90	3.5	1285	2	US-08-453-141-2	Sequence 2, Appli
39	90	3.5	1285	3	US-08-293-314-2	Sequence 2, Appli
40	90	3.5	2958	4	US-08-894-344C-2	Sequence 2, Appli
41	89.5	3.5	969	1	US-08-365-689-3	Sequence 3, Appli
42	89.5	3.5	969	1	US-07-747-781-3	Sequence 3, Appli
43	89.5	3.5	969	1	US-08-145-138A-3	Sequence 3, Appli
44	89.5	3.5	969	5	PCT-US92-06888-3	Sequence 3, Appli
45	89.5	3.5	969	5	PCT-US93-03027-4	Sequence 4, Appli

ALIGNMENTS

RESULT 1	US-09-134-001C-5355	Sequence 5355, Application US/09134001C	Patent No. 6380370
GENERAL INFORMATION:	APPLICANT: Lynn Doucette-Stamm et al	TITLE OF INVENTION: NOCLETIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCC	FILE REFERENCE: GTC-007
CURRENT FILING DATE: 1998-08-13	PRIOR APPLICATION NUMBER: US 60/064,964	PRIOR FILING DATE: 1997-11-08	PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14	NUMBER OF SEQ ID NOS: 5674	SEQ ID NO 5355	LENGTH: 362
TYPE: PRT	ORGANISM: Staphylococcus epidermidis	US-09-134-001C-5355	Query Match
Best Local Similarity 19.8%, Pred. No. 3e-06;	Matches 94; Conservative 72; Mismatches 167; Indels 141; Gaps 20;	Score 142.5; DB 4; Length 362;	Best Local Similarity 19.8%, Pred. No. 3e-06;
Matches 94; Conservative 72; Mismatches 167; Indels 141; Gaps 20;	Score 142.5; DB 4; Length 362;	Best Local Similarity 19.8%, Pred. No. 3e-06;	Matches 94; Conservative 72; Mismatches 167; Indels 141; Gaps 20;
3 VVASAPGKVLMTGGLVLEKPNAGLVLTSTNARFVAIVKPEVKEPESWPKMTQVTKLTS 62	6 IOVKAPGKIYINGEVAIVPEGKSTILIANRVTVITRESNK-----VGSITS 54	63 POLRESM-YKLSLHNTLOSASDSRNPFEVHAIOVAIAAHLATERKESLKLILQ 121	55 KTLHEPVKFDNENRIEISDVQAAKQ-----LKYVVAIEV----- 91
122 GIDITILSNDFYSTRNOIESAGLPLTPESLCTLAPFASITFNAESGANSKPEVATG 181	92 -----PEQVYRSCNMMLKHFHL-----TIDSNLADNSG-----QRYG 123	132 LGSSAAMTAVVAALHLHYGVVDSDPCKEKGFGSGSDIDVHIMIQTSCLAQGVSGGF 241	134 LGSSAALVSVYKALNEFYG-LELSNL-----YIKVLAVINMKLQSLSSCG- 169
242 DVSCAVY-GSORRYAFSPSEVLFAOVAATGLPLNEVIGTILKGKDNKRTF-FSLPLMN 299	170 DIASVYSGWLAISTFDHDKW-KQMEETS-----VNDVLEKNPGLHIEQLAPENME 222	300 LFLGPGSGSSTPSMGAVKWKQMSDPEKARENQNLSDANLELEYTLNDLSKLAKDH 359	223 VLIGWTGSPASS-PLHVEVVKRLK-SDP-----SFY 251

QY 360 DYLRVIRKSCSVLTSEKWLHATEPIINEAIKELLEAREMLRI-RILMOMGEAAVPI 418
Db 252 GDELDSHAC-----VESLIQAFKTNKICGVQKIRINRIRRIISDMOEASVEI 299
QY 419 EPESQQLDSTMSABGVLLAGVPAGFAIRAITIGDSGTIKLTQWSSHNVL 472
Db 300 ETDKLKLCVGEKHGA--SKTSGAG-----GDCGITTINKVIDKNII 342

RESULT 2

5349058-2
Patent No. 5349058
APPLICANT: TANAKA, RICHARD D.; RICCI, BEVERLY S.
MOSLEY, STEPHEN T.
TITLE OF INVENTION: NUCLEIC ACID ENCODING HUMAN
MEVALONATE KINASE
NUMBER OF SEQUENCES: 2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/475,577
FILING DATE: 06-FEB-1990
SEQ ID NO: 2
LENGTH: 396
5349058-2

Query Match 4.5%; Score 117; DB 6; Length 396;
Best Local Similarity 20.2%; Pred. No. 0.0017;
Matches 104; Conservative 74; Mismatches 183; Indels 154; Gaps 22;

QY 3 VVASAPGKVLMTGGYLVLEKPNAGLVSTNARFAIYKPIINEEVKPPSMAMKMTDVLTS 62
Db 6 LTVSAPGKVLHGEHAYVH-GKVALAVSLNRTFLRLQPHSN-----GKVDLSL 53
QY 63 POLSRESMYKLSLNHLTLQSVASDSRNPFVEHAIOYAIAAHLATEKDESLHKLQ 122
Db 54 PNIGIKRAMOVA-----RLQSLDTS-----FLEOGD-----VTPTSEQVEKLR--- 92
QY 123 LDITILGSDNPFYRNQIESAGLP---LTPESIGTLAPPASITITNAESGANSKPEVA- 178
Db 93 -----EVAGLPDDCAVTERLAIVLA-FYLYLISICKORALPSLDIV 133
QY 179 -----KTGLSSNAAMTAVAAALHLYGVNLSDPCKEG---KFGCSLDIVHMTAQT 228
Db 134 WSELPPAGIGSSNAYVCCLAALITY--CEETPNPLKDDCVAKWKKEDELEINKAFO 191
QY 229 SHCIAQKVGSGPDVSCAVYGSQRYVRESPEVLSFAQVAVTGPLNEVIGITILGKWDNK 288
Db 192 GERIHGN-PSGVNNAVSTWGG-----ALRYHOGKISSLSKSPALQILL----- 234
QY 289 RTEISLPLMNLFLGEGSGSSTPSMVGAVKKWQMSDPEKARENMONLSDALEETKL 348
Db 235 -TNTKVP-----RMTALVAGVRNRLKFPETIVAPLTLTSDAISLECERVL 279
QY 349 NDLSKLAKDHWYLRVIRKSCSVLTSEKWLHATEPIINEAIKELLEAREMLRIRILMR 408
Db 280 GEMGE-----APAEQVLYLEELIMNQHHLN----- 306
QY 409 OMGEAASVPIEPESQQLDSTMSABGVLLAGVPAGFAIRAITIGDSGTIKLTQWSS 468
Db 307 -----ALGVGHASLDQLCVLT-RARG-LHSKILGAGGGCC-GITLLKGLQEPEVAT 356
QY 469 HNVIALLVREDPHGVCELSGDPRTTCTTSCVSSIH 503
Db 357 KQALTSGCFD-----CLE-----TSIGAPGVSIH 380

RESULT 3

US-09-306-595C-8
Sequence 8, Application US/09306595C
Patent No. 6284506
GENERAL INFORMATION:
APPLICANT: HOSHINO, Tatsuo
APPLICANT: OJIMA, Kazuyuki
APPLICANT: SETOGUCHI, Yutaka

TITLE OF INVENTION: ISOPRENOLID PRODUCTION
FILE REFERENCE: ISOPRENOLID PRODUCTION
CURRENT APPLICATION NUMBER: US/09/306,595C
CURRENT FILING DATE: 1999-05-06
PRIOR APPLICATION NUMBER: 98108210
PRIOR FILING DATE: 1998-05-06
NUMBER OF SEQ ID NOS: 43
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 8
LENGTH: 432
TYPE: PRT
ORGANISM: Phaffia rhodozyma
US-09-306-595C-8

Query Match 4.5%; Score 115; DB 4; Length 432;
Best Local Similarity 21.6%; Pred. No. 0.0032;
Matches 103; Conservative 64; Mismatches 190; Indels 120; Gaps 19;

QY 3 VVASAPGKVLMTGGYLVLEKPNAGLVSTNARFAIYKPIINEEVKPPSMAMKMTDVLTS 62
Db 4 LTVSAPGKVLHGEHAYVH-GKVALAVSLNRTFLRLQPHSN-----GKVDLSL 53
QY 63 POLSRESMYKLSLNHLTLQSVASDSRNPFVEHAIOYAIAAHLATEKDESLHKLQ 121
Db 49 SLSSTNVI-TTSLDNLNFTQSWPVDS-----LPWSLAPDW--TEASTPESICPTLLA 97
QY 122 GLD-TTILGSDN-----FYSRNQIESAGLPPLPESIGTLAPPASITITNAESNGA 171
Db 98 EIERIAGGNGGNGEREKAVATMAFLYLVLLSKGKPSPEFL-----TASALPMDGA 148
QY 172 NSKPEVAKTGLSSNAAMTAVAAALHLYGVNLSDPCKEGKGCSDLYIHMTAOTSCH 231
Db 149 -----GLGSSALSTLALVFLHFSHLSPTTGRESTTPTADTEVIDRWAFLEK 199
QY 232 LAQKVGSGPDVSCAVYGSQRYVRESPEVLSFAQVAVTGPLNEVIGITILGKWDNKRT 291
Db 200 VTHGN-PSGIDNAVSTWGG-----AVAKRKRIEKGQ-E 230
QY 292 FSLPPLMNL-----FLGEPGSGSSTPSMVGAVKKWQMSDEK-----ARENMON 336
Db 231 GGMFAIKSFTSIRPLTDSRIGRDTSLVAGVANRLIQEDEVIVPLLEAIOQTADAEAIRC 280
QY 337 LSDANLELEKRLNDLSLAKDHWYLRVIRKSCSVLTSEKWLHATEPIINEAIKELLEA 356
Db 291 LKDSMERAVMIRLOMLVSEN-----HAHLAAGVSHPSIEEIIIR-IGAD 335
QY 397 REAMLRIIRIMROGEAASVPIEPESQQLDSTMSABGVLLAGVPAG-GF 447
Db 336 KPPELRIKILGAGGGCAVTLVDPDEFTEFLQALMETLVGSSFAPIYARVGGSGVGF 392

RESULT 4

US-08-038-682-4
Sequence 4, Application US/08038682
Patent No. 5549897
GENERAL INFORMATION:
APPLICANT: BARENKAMP, STEPHEN J
APPLICANT: ST. GEME III, JOSEPH W
TITLE OF INVENTION: HIGH MOLECULAR WEIGHT SURFACE PROTEINS
TITLE OF INVENTION: OF NON-TYPEABLE HAEMOPHILUS
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Shoemaker and Matlare, Ltd
STREET: 2001 Jefferson Davis Hwy., 1203 Crystal Plaza
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202-0286
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/038,682
FILING DATE: 16-MAR-1993
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: BERKSTRESSER, JERRY W
REGISTRATION NUMBER: 22,651
REFERENCE/DOCKET NUMBER: 1038-293
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 415-0810
TELEFAX: (703) 415-0813
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1477 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-038-682-4

Query Match 4.2%; Score 108.5; DB 1; Length 1477;
Best Local Similarity 20.3%; Pred. No. 0.13; Mismatches 162; Indels 119; Gaps 21;
Matches 89; Conservative 68;

QY 11 VLMTGGLVLEKPNPNA-----GLVLTNARFYAIKPINEEYKPSMAWKMT-----56
DB 139 ILDSNGGVFLNPGITIGKDAIINTNGFTASTLDSINENIKARNFTEQYKDKALAEIV 198
QY 57 -----DVKLTSQLSRESMYKL--SLNHLTLQSVASASDRNPVEHAIOYAIA 102
DB 199 NHGLITVGKDSVNLIGKVKNEGIVSYNGSISILAGOKITISDIINP-----TITYSIA 254
QY 103 AAHLATEKDKESLKLKLLQGLDITILGSDNYSYRNOIESAGLPLTPESLGLAPFASIT 162
DB 255 -----APENEAVNLGDIFRAKGNINVRRA-----TIRNOKLSADSVSKDKSGNIV--LSAK 304
QY 163 FNAESNGA-NSKPEVAKTG-----LGSSAAMTAVVAALL-----HYLGVDLSDPCKE 211
DB 305 EGEAIEIGVISAQNOQAKGKMLTGKVLTKGAVIDLSGKEGGEYLG-----GDERGE 360
QY 212 GKFGCSDLDVTHMTAQ-----SHCLAQKVGSGFDVSCAVYGSORRYRSEPEVLSPAOV 266
DB 361 GKNG-----IOLAKKTSLEKSTINVSKEKEGGFAI--VWGDIALIDGINAOGSGDI 411
QY 267 AVTG-----LPLNEVIGTILKGM-----DNKRTESLPLIMLFLGEPGSGSS 311
DB 412 AKTGGEVETSGHDLFIKDNALVDA---KEWLDPDNVSINADPLFNNTGTGINDPEPTGTG 468
QY 312 TPSNVAVKKQMSDPEKARENMONLSDANLELETKLND--LSKLAKDHWV-----YL 363
DB 469 -----EASDPKK-----NSELKTTLLNTTISYVLKNAAMTNITASKRL 506
QY 364 RVIKSCSVLTSEKWLHA 381
DB 507 TVNSSINIGNSHLIHS 524

RESULT 5
US-08-302-832-4
Sequence 4, Application US/08302832
Patent No. 5603938
GENERAL INFORMATION:
APPLICANT: Barenkamp, Stephen J
TITLE OF INVENTION: High Molecular Weight Surface Proteins
TITLE OF INVENTION: of No. 5603938-Typeable Haemophilus
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Shoemaker and Matlare, Ltd.
STREET: 2001 Jefferson Davis Hwy., 1203 Crystal Plaza
CITY: Arlington
STATE: Virginia

COUNTRY: U.S.A.
ZIP: 22202-0286
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/302,832
FILING DATE: 16-SEP-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9205704.1
FILING DATE: 16-MAR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US pct/us93/02166
FILING DATE: 16-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Berkstresser, Jerry W
REGISTRATION NUMBER: 22,651
REFERENCE/DOCKET NUMBER: 1038-404
TELEPHONE: (703) 415-0810
TELEFAX: (703) 415-0813
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1477 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-302-832-4

Query Match 4.2%; Score 108.5; DB 1; Length 1477;
Best Local Similarity 20.3%; Pred. No. 0.13; Mismatches 162; Indels 119; Gaps 21;
Matches 89; Conservative 68;

QY 11 VLMTGGLVLEKPNPNA-----GLVLTNARFYAIKPINEEYKPSMAWKMT-----56
DB 139 ILDSNGGVFLNPGITIGKDAIINTNGFTASTLDSINENIKARNFTEQYKDKALAEIV 198
QY 57 -----DVKLTSQLSRESMYKL--SLNHLTLQSVASASDRNPVEHAIOYAIA 102
DB 199 NHGLITVGKDSVNLIGKVKNEGIVSYNGSISILAGOKITISDIINP-----TITYSIA 254
QY 103 AAHLATEKDKESLKLKLLQGLDITILGSDNYSYRNOIESAGLPLTPESLGLAPFASIT 162
DB 255 -----APENEAVNLGDIFRAKGNINVRRA-----TIRNOKLSADSVSKDKSGNIV--LSAK 304
QY 163 FNAESNGA-NSKPEVAKTG-----LGSSAAMTAVVAALL-----HYLGVDLSDPCKE 211
DB 305 EGEAIEIGVISAQNOQAKGKMLTGKVLTKGAVIDLSGKEGGEYLG-----GDERGE 360
QY 212 GKFGCSDLDVTHMTAQ-----SHCLAQKVGSGFDVSCAVYGSORRYRSEPEVLSPAOV 266
DB 361 GKNG-----IOLAKKTSLEKSTINVSKEKEGGFAI--VWGDIALIDGINAOGSGDI 411
QY 267 AVTG-----LPLNEVIGTILKGM-----DNKRTESLPLIMLFLGEPGSGSS 311
DB 412 AKTGGEVETSGHDLFIKDNALVDA---KEWLDPDNVSINADPLFNNTGTGINDPEPTGTG 468
QY 312 TPSNVAVKKQMSDPEKARENMONLSDANLELETKLND--LSKLAKDHWV-----YL 363
DB 469 -----EASDPKK-----NSELKTTLLNTTISYVLKNAAMTNITASKRL 506
QY 364 RVIKSCSVLTSEKWLHA 381
DB 507 TVNSSINIGNSHLIHS 524

RESULT 6
US-08-530-198-4
Sequence 4, Application US/08530198
Patent No. 5869065


```

0Y      212 GFFGSGDDIVHIMIAQT-----SHLQAGKVGSGDVSACAVYGSGQRVYFSPSEVLSPAAY 266
Db      361 GRNG-----IOLAKTISLEKGSTTNVSGKEKGFAI---WGDIALIDGNINNAOGSGDI 411
0Y      267 AVTG-----LPINEVIGITILKGKW----DNKRTEFSIPPLMNLFLGEPSGGSS 311
Db      412 AKTGGFVEFTSGHDLFIKDNAIVDA---KEMLLDPFNVSINADPDLFNNNTGINIDEPTGTG 468
0Y      312 TPSMWGAVKWKQMOSDEPKARENMONLSDAELETKLAND--LSKLAKDHWDY-----YL 363
Db      469 -----EASDPKK-----NSBLKTTLTFWTFTISNYLKNAWMTNNTIASRKL 506
0Y      364 RYKSCSVLTSEKMYLHA 381
Db      507 TVNSSINIGSNSHLILHS 524

RESULT 8
US-08-728-470-4
Sequence 4, Application US/08728470
Patent No. 5928651
GENERAL INFORMATION:
APPLICANT: Barenkamp, Stephen J
TITLE OF INVENTION: High Molecular Weight Surface Proteins
TITLE OF INVENTION: of No. 5928651-Typeable Haemophilus
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Shoemaker and Mattare, Ltd.
STREET: 2001 Jefferson Davis Hwy., 1203 Crystal Plaza
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202-0286
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/728.470
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/302.832
FILING DATE: 16-MAR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US PCT/US93/02166
FILING DATE: 16-MAR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9205704.1
FILING DATE: 16-MAR-1992
ATTORNEY/AGENT INFORMATION:
NAME: Berkstresser, Jerry W
REGISTRATION NUMBER: 22,651
REFERENCE/DOCKET NUMBER: 1038-633
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 415-0810
TELEFAX: (703) 415-0813
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1477 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-728-470-4

Query Match          4.2%; Score 108.5; DB 2; Length 1477;
Best Local Similarity 20.3%; Pred. No. 0.13;
Matches    89; Conservative   68; Mismatches 162; Indels 119; Gaps 21.
```

```

Db 139 ILDSNGQVFLNPNIGITGKALLINTNGFTASTJLDJENENIKANFPEFGTKOKALAEIV 198
OY 57 -----DKVLTSPQLSRESMYKL---SLNHLTLOSVSASDSRNPVEHAIOYAIA 1020
Db 199 NHGLITVKGDSVNLIGKVKNEGVISVNGGISLLAGOKTITSDIINP-----TIYYSIA 254
OY 103 AAHLATEKDKSILKRLLOGDITITLGSNDPYSVRNQIESAGLP/LPESIGTLAPFSTI 162
Db 235 ---APEAEVNLIDIFKKGGINIVRA---TIRNGCKLSADSVSKDKGNIV--LSAK 3040
OY 163 FMAASNA--NSKPEVATKG---LGSAAATTAVVAALL-----HYLGVDLSDPCKE 2111
Db 305 EGEAFETGIVISAQNOQAKGKMLMTGDKVTLKTAVIDLSKKEGGEYTLG---GDRGE 3600
OY 212 GKFGCSDLDVIMHIAQT-----SHCLAQKGVSGFDVSCAVYGSQRYVRESPEVLSFAOV 266
Db 361 GKNG-----IQLAKKTSLEKGTSTINVSKEKGFAL---VMDGIALIDGINAAGSGDI 4111
OY 267 AVTG-----LPANEYIGITLGGKW---DNKRTEPSLPPLNNL/LGEGSGSS 3111
Db 412 AKTGEFVETSGHDLFIKNAIVDA---KEWLDPEDNVSIINAEDPLFNNTGINDPEFGTGTG 468
OY 312 TPSWGAVKAKMQMSDEPKARENMQLSDANLEETKILND--LSLAKDHDV-----YL 363
Db 469 -----EASDPK-----NSELKTTTLTNTTISYLNKAMTAMTATSRKL 506
OY 364 RVKSCSVLTSEKWVLA 381
Db 507 TVNSSINIGNSHLILHS 524

RESULT 9
US-08-617-697-4
: Sequence 4, Application US/08617697
: Patent No. 5977336
: GENERAL INFORMATION:
: APPLICANT: Barenkamp, Stephen J
: TITLE OF INVENTION: High Molecular Weight Surface Proteins
: TITLE OF INVENTION: of No. 5977336-Typeable Haemophilus
: NUMBER OF SEQUENCES: 11
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Shoemaker and Mattare, Ltd.
: STREET: 2001 Jefferson Davis Hwy., 1203 Crystal Plaza
: STREET: Bldg. 1
: CITY: Arlington
: STATE: Virginia
: COUNTRY: U.S.A.
: ZIP: 22202-0286
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent In Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/617,697
: FILING DATE: 01-APR-1996
: CLASSIFICATION: 424
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/302,832
: FILING DATE: 05-OCT-1994
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US PCT/US93/02166
: FILING DATE: 16-MAR-1993
: ATTORNEY/AGENT INFORMATION:
: NAME: Berkstresser, Jerry W
: REGISTRATION NUMBER: 22,651
: REFERENCE/DOCKET NUMBER: 1038-557
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (703) 415-0810
: TELEFAX: (703) 415-0813
: INFORMATION FOR SEQ ID NO: 4:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1477 amino acids
:

```


SEQ ID NO 71
 : LENGTH: 1477
 : TYPE: PRT
 : ORGANISM: Haemophilus influenzae
 US-09-206-942-71

Query Match 4.2%; Score 108.5; DB 4; Length 1477;
 Best Local Similarity 20.5%; Pred. No. 0.13;
 Matches 90; Conservative 70; Mismatches 157; Indels 123; Gaps 21;

QY 11 VLMITGVLEKPN-----GLVSTNARFYAIKPIINEEYKPPSMARKMT-----56
 DB 139 ILDSNGCVFLPNIGITIGKDAIINTNGFTASTLDISNENIKANFTFEQTKALAEIV 198
 QY 57 -----DVKLTSPOLSRSMYKL---SLNHLTLQASASDSRRPFENALQYAI 102
 DB 199 NHGLITWGKDSVNLIGKAKNEGVISVNGSISLAGOKTITSDIINP---TITYSIA 254
 QY 103 AAHLATEKDESLHKLLOGIDITILGSDNFYSYRNOIESAGLPTEPSLCTLPASFAT 162
 DB 255 -----APENEAVNLGDIKAKGNINVRAA---TIRNOGKLADSVSKDNGNIV--LSAK 304
 QY 163 FNAESNGA-NSKPEVAKTG---LGSSAANTYVAALL-----HYGVYDLSDPCKE 211
 DB 305 EGEAEIGGVISAQNOAKGKLMITGDKVTLTGAVIDLSEKGEGETYLG---GDERGE 360
 QY 212 GKFCGSDLDVIMIAOTSHCLAQGVSGFDVS-----CAYYGSQRYVRFSPVLSFA 264
 DB 361 GKNG-----IQLAKTSL-----EKGSTINVSKEKKGRAIVMGDIALIDGNINAAQSG 409
 QY 265 QVAVTG-----LPLNEVIGTILGKW---DNKRTEFSLPPLMLFLGEPGSGG 309
 DB 410 DIARTGFEVETSGHYLSIDSNAIVKT---KEMLDPPDVTLEADPLNNNGINDEPPTG 466
 QY 310 SSTSMGAVKAKKQMSPEKAREMQUISDANLETERKLN--LSKLAKHMDV-----361
 DB 467 TG-----EASDPK-----NSELKTTLTNTTISNLYKNAKMTNNTASR 504
 QY 362 YLRVYKSCVLTSEKWLHA 381
 DB 505 KLTIVNSSINIGSNHLIHS 524

RESULT 12
 US-09-651-200-4
 : Sequence 4, Application US/09651200
 : Patent No. 6429303
 : GENERAL INFORMATION:
 : APPLICANT: Green et al
 : TITLE OF INVENTION: Polynucleotides Encoding Members of the Human B
 : TITLE OF INVENTION: Lymphocyte Activation Antigen B-7 Family and
 : TITLE OF INVENTION: Polypeptides Encoded Thereby
 : FILE REFERENCE: 15966-562 (CURA-62)
 : CURRENT APPLICATION NUMBER: US/09/651.200
 : CURRENT FILING DATE: 2000-08-30
 : PRIOR APPLICATION NUMBER: 60/152383
 : PRIOR FILING DATE: 1999-09-03
 : PRIOR APPLICATION NUMBER: 60/172909
 : PRIOR FILING DATE: 1999-12-21
 : PRIOR APPLICATION NUMBER: 60/183578
 : PRIOR FILING DATE: 2000-02-18
 : NUMBER OF SEQ ID NOS: 25
 : SOFTWARE: Patent In Ver. 2.0
 : SEQ ID NO 4
 : LENGTH: 441
 : TYPE: PRT
 : ORGANISM: Homo sapiens
 US-09-651-200-4

Query Match 4.0%; Score 104.5; DB 4; Length 441;
 Best Local Similarity 22.3%; Pred. No. 0.042;
 Matches 87; Conservative 45; Mismatches 112; Indels 147; Gaps 22;

QY 209 CKEKFGCSDLDVIMIAOTSHCLAQKV---GSGFDVSCAVYGSQRYVRFSPVLSFAV 266
 DB 39 CRQAPGKSLLE-----AQHDPGAQOGPAARGHGVITICSSY--QGY-----PEAEVFWQD 87
 QY 267 AVTGLPLNEVIGTILGKWKDKKRTFSLPPLMLFLGEPG-----SGGS 310
 DB 88 G-QGVPLT---GNVTTSSMANEGCLFDVHSILRVVLGANGTYSCLVRNPVLQODAHSSVT 143
 QY 311 STP---SMGAVKAKKQMSPEKAREMQUISDANLE-----TKLNDLSKLANQHW 359
 DB 144 ITPQSRPTGAIVEVQVPEDPVALVG---TDATLHCSFSPPEPFSLTQNLIL-----W 192
 QY 360 DVYLRVYKSCVLTSEKWLHA-TEPINE-----AIKELLEAREAMLRIR-----404
 DB 193 Q-----LDTKOLVHSFTFEGRQGSAYANRRALFPDLLAOGASLRLOQRYRAD 241
 QY 405 -----ILMRQGEA-----SVPIEPESOTQLDST-----430
 DB 242 EGSFTCEFSIRDFGSAVSLQVAAPYSKPSMTLEPNKDLRPEDVTITCSSYRGYPEAEV 301
 QY 431 --MAEGVLLG-----VPGAGGFD--AIFAITLGDSSGKTLQAMSHVALLVR-- 477
 DB 302 FWDQGGVPLTGNVTTSSMANEGCLFDVHSILRVVLGANGT-----YSCLVANP 350
 QY 478 --EDPHGVCLSESDPRT---TCITSGVS 500
 DB 351 VLQGDHGSVITIGQPMFPEPALMVTYGLS 381

RESULT 13
 US-08-445-135-2
 : Sequence 2, Application US/08445135
 : Patent No. 5658789
 : GENERAL INFORMATION:
 : APPLICANT: Quaranca, Vito
 : APPLICANT: Hormla, Markella
 : TITLE OF INVENTION: Promotion of Epithelial Cell Adhesion
 : TITLE OF INVENTION: and Hemidesmosome Assembly by a Laminin-like Molecule
 : NUMBER OF SEQUENCES: 7
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Knobbe, Martens, Olson and Bear
 : STREET: 620 Newport Center Drive, Sixteenth Floor
 : CITY: Newport Beach
 : STATE: CA
 : COUNTRY: USA
 : ZIP: 92660
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: Patent In Release #1.0, Version #1.25
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/08/445.135
 : FILING DATE:
 : CLASSIFICATION: 435
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: US 08/151,134
 : FILING DATE: 12-NOV-1993
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Israelien, Ned A.
 : REGISTRATION NUMBER: 29,655
 : REFERENCE/DOCKET NUMBER: DESMOS.002CPE2
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: (619) 235-8550
 : TELEFAX: (619) 235-0176
 : INFORMATION FOR SEQ ID NO: 2:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 770 amino acids
 : TYPE: amino acid
 : STRANDEDNESS: single
 : TOPOLOGY: linear
 : MOLECULE TYPE: protein
 : HYPOTHEICAL: NO

ANTI-SENSE: NO
FRAGMENT TYPE: Internal
IMMEDIATE SOURCE:
CLONE: 150 KD
US-08-445-135-2

Query Match 4.08; Score 104.5; DB 1; Length 770;
Best Local Similarity 19.18; Pred. No. 0.11;
Matches 98; Conservative 65; Mismatches 196; Indels 153; Gaps 17;

QY 67 RESM--YKLSLNLH--TIGSVASDSNRPVFAIOLYAIAMAHATKDESLHKLLOG 122
DB 125 RDSLNDTEAKLQDRSVYQAAAGKATGLNHNESVLAIQ-RQKEMDSLKYITLH 183
QY 123 L--DITILGNDYFYRNQIESAGLPETPEISLGTLLAPFASITFNAESGANSKEPVAK 179
DB 184 LAVADASLQTNLSLQMDTSQK-----EYESL-----AALNGARQELNDQY 226
QY 180 TGLGSSAMTTAVVALLHLGLVVDLSDPCKEKGCGSDLDVIMIAQTSCHL----- 232
DB 227 RELSRSGKAPLYAEAEKHAQSLOELAKOLEIKRNTSGESVRCVDAATAVESIIMAI 286
QY 233 -----AGKYGSGPDYSCAVYSGRYRFSPEVLSFAQVAVTGLP-----LN 274
DB 287 RAEDDAAGKADSASE-----SAFOTVAKEDLPRAKTLSDSEELLN 328
QY 275 EVIGTILKGMKDKREFSLPPLMN-----LFLGEP 305
DB 329 E-----AKMTKRRLQOEINPALNSLOTLKTVSVQKDLLANVAVRNDLNGIORGDI 381
QY 306 GSGGSSTPSVV---GAVKKQMSDPKAREWQNL-----SDANLETRKNDLSK 353
DB 382 DSVVSGAKSVRRKANGITSEVLDSAPSRRIMEGLRTATCTRHEDPKALIMANSVKK 441
QY 354 LAKDHMDVYLRVJKSCSVLSEKVVHATEPINEAI-----IKEL-EAREAM 400
DB 442 LTKKLPLPVKI-----ESINQOLPLIGNISNVRIEELITIOARDA 484
QY 401 LRIRILRMQGEASVPIEPESOTQLDSTMSAEGVLLAGVPGAGGFDAITLGDST 460
DB 485 NKVAIPRFKNGKSGVEYRLNDELDTKGYTSLFLQRPDLRENGTJEDMFVYLGKND- 543
QY 461 KLGQAMSHNVALLVREDPHGVCLESGDPT 492
DB 544 -----ASKDYIGIGCGRPADVCLQGGPRS 569

RESULT 14
US-09-651-200-6
Sequence 6, Application US/09651200
Patent No. 6429303

GENERAL INFORMATION:
APPLICANT: Green et al
TITLE OF INVENTION: Polynucleotides Encoding Members of the Human B
TITLE OF INVENTION: Lymphocyte Activation Antigen B-7 Family and
TITLE OF INVENTION: Polypeptides Encoded Thereby
FILE REFERENCE: 15966-562 (CURA-62)
CURRENT APPLICATION NUMBER: US/09/651,200
CURRENT FILING DATE: 2000-08-30
PRIOR APPLICATION NUMBER: 60/152383
PRIOR FILING DATE: 1999-09-03
PRIOR APPLICATION NUMBER: 60/172909
PRIOR FILING DATE: 1999-12-21
PRIOR APPLICATION NUMBER: 60/183578
PRIOR FILING DATE: 2000-02-18
NUMBER OF SEQ ID NOS: 25
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 6
LENGTH: 534
TYPE: PRT
ORGANISM: Homo sapiens
US-09-651-200-6

Query Match 3.98; Score 101.5; DB 4; Length 534;
Best Local Similarity 22.88; Pred. No. 0.12;
Matches 97; Conservative 48; Mismatches 136; Indels 145; Gaps 23;

QY 183 GSSAAMTTAVVALLH-----LGVVDLSDPCKEKGCG-----SDLDVYHMI 225
DB 86 GSAVANRTALFPDLLAOGNASLRQRYRVAD---EGSFTCFVSIRDEGSAVSLQVAPY 142
QY 226 AOTSCHLAQK---VSGSFPVSCAVYSGRYRFSPEVLSFAQVAVTGLPINEYITL 282
DB 143 SKPSMTLEPKDLRPGDTVITCSSY--QGY-----PEAEVFMQDG--QGVPLT--GNVTT 192
QY 283 GKMDNKRTEFSPLPLNLFGEPC-----SGSSSP--SMGAVKKQW 324
DB 193 SQMANEGELPDVHSIILRVVANGYSCILVRNPVLAQDAHSSVTITPQRSPTGAVEQVP 252
QY 325 SDEPKAREMNQNSDANLELETKLNDLSLAKDHMDVYLRVISCVSLVSEKVVLAH-TE 383
DB 253 EDVVALVG---TDATLRCSFSPBEPGSLAQ-----LNLWQ---LDTTKOLVHSFTE 299
QY 384 PINE-----AIKELREAREAMLRIR-----ILMRQGEAA----- 414
DB 300 GROGSAVANRTALFPDLLAOGNASLRQRYRVADSGSFTCFVSIRDEGSAVSLQVAPY 359
QY 415 -----SVPIEPESOTQLDST-----MSAEVLLAG-----VEGAGG 446
DB 360 YSKPSMTLEPKDLRPGDTVITCSSYRGYPBAEVFMQDGVPVLTGNVTTSQMANBOGL 419
QY 447 FD--AIFAITLGDSTGLTQAMSHNVALLVR-----EDPHGVCLESGDPT-----TC 494
DB 420 FDVHSAVLRVVLGANGT-----YSLVRNPVLAQDAHSSVTITGQPMFFPEALM 468
QY 495 ITSGVS 500
DB 469 VTFVGLS 474

RESULT 15
US-09-651-200-24
Sequence 24, Application US/09651200
Patent No. 6429303

GENERAL INFORMATION:
APPLICANT: Green et al
TITLE OF INVENTION: Polynucleotides Encoding Members of the Human B
TITLE OF INVENTION: Lymphocyte Activation Antigen B-7 Family and
TITLE OF INVENTION: Polypeptides Encoded Thereby
FILE REFERENCE: 15966-562 (CURA-62)
CURRENT APPLICATION NUMBER: US/09/651,200
CURRENT FILING DATE: 2000-08-30
PRIOR APPLICATION NUMBER: 60/152383
PRIOR FILING DATE: 1999-09-03
PRIOR APPLICATION NUMBER: 60/172909
PRIOR FILING DATE: 1999-12-21
PRIOR APPLICATION NUMBER: 60/183578
PRIOR FILING DATE: 2000-02-18
NUMBER OF SEQ ID NOS: 25
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 24
LENGTH: 534
TYPE: PRT
ORGANISM: Unknown
FEATURE: Description of Unknown Organism: Sequence
OTHER INFORMATION: mz5020.protein from Figure 4.
US-09-651-200-24

Query Match 3.98; Score 101.5; DB 4; Length 534;
Best Local Similarity 22.88; Pred. No. 0.12;
Matches 97; Conservative 48; Mismatches 136; Indels 145; Gaps 23;
QY 183 GSSAAMTTAVVALLH-----LGVVDLSDPCKEKGCG-----SDLDVYHMI 225
DB 86 GSAVANRTALFPDLLAOGNASLRQRYRVAD---EGSFTCFVSIRDEGSAVSLQVAPY 142


```

0Y 226 AOTSICLAQG---VGSFDFSCAVVGSQRYRRSEPEVLSFAQVAAVYGLPIANEVICTILK 282
Db 143 SKPSTLTLPNDLRGDTVTITGSSY--QGT---PEAEVFWODG--OGVLT---GNVTT 1922
0Y 283 GKMDNRKRFESLPLMLNLFGEPP-----SGGSSPP--SMGAVKMKQM 324
Db 193 SOMANEGLFVHSLIRVHSLANGANTYSCLYRNPVLOQDAHSSVTITPQRSPTCAVEYQP 252
0Y 325 SDPEKARENMONLS DANLEETKLNDSIKLAKDHMDVYLLRVIKSCSVLTSEKVVLAH-TE 383
Db 253 EDPVVALVG---TDATLRCSFSEPEGFLSAO---LMLIMQ--LFTDKQLVHSPTE 209
0Y 384 PINE-----AIKKELEKREMLAIR-----ILMROMGAA----- 414
Db 300 GRDGSAYANFTALEPDLLAGNNSLRLOVRVRADEGSCFTCFYSIRPFGSAAYSLQVAP 359
0Y 415 ---SVPIPEESOTQLDST-----MSAEGVILAG-----VPGAG 446
Db 360 YSKPSMTLEPRKKDLRPGDTVITTCSSYRGYPEAEVFWODGQGYPLTGNVTTSQMANOGL 419
0Y 447 FD--AIFAITLGDGCTKLQAMSSHNVALIVR-----EDPHGVCLESGDPT-----TC 494
Db 420 FDIHVSLIRVVLGANGT-----YSCLYRNPVLOQDAHGSVTTTGPMPFPPEALM 468
0Y 495 ITSGVS 500
Db 469 VTVGLS 474

```

Search completed: April 26, 2003, 12:57:54
Job time : 22 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4_p5-4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 26, 2003, 12:57:10 ; Search time 19 Seconds
(without alignments)
2129.766 Million cell updates/sec

Title: US-09-988-863a-2

Perfect score: 2581
Sequence: 1 MAVVASAPGKVLMTGTYLVL.....ESGDPRTCTTSGVSSIHLE 505

Scoring table: -BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 301932 seqs, 80129803 residues

Total number of hits satisfying chosen parameters: 301932

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

Published_Applications_AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2581	100.0	505	10	US-09-988-863a-2
2	575	22.3	179	10	US-09-909-745-22
3	453.5	17.6	451	10	US-09-909-745-24
4	216	8.4	67	10	US-09-909-745-20
5	204	7.9	54	10	US-09-909-745-16
6	204	7.9	54	10	US-09-909-745-18
7	178.5	6.9	368	10	US-09-815-242-10656
8	135.5	5.2	358	10	US-09-815-242-12866
9	133.5	5.2	345	10	US-09-815-242-5854
10	126	4.9	386	10	US-09-909-745-26
11	123	4.8	378	10	US-09-909-745-23
12	115	4.5	432	9	US-09-925-388-8
13	108.5	4.2	1477	9	US-10-092-880-4
14	101.5	3.9	534	9	US-10-077-023-7
15	101.5	3.9	534	10	US-09-875-338-7
16	101	3.9	1025	8	US-08-910-386a-7
17	100	3.9	335	10	US-09-815-242-13670
18	100	3.9	698	9	US-10-077-023-9
19	100	3.9	698	10	US-09-875-338-9

20	97	3.8	863	10	US-09-946-239-11	Sequence 11, Appl
21	96	3.7	657	10	US-09-815-242-13436	Sequence 13436, A
22	96	3.7	854	9	US-09-770-107-2	Sequence 2, Appl1
23	95.5	3.7	1599	9	US-10-092-880-9	Sequence 9, Appl1
24	95	3.7	492	9	US-09-738-626-5963	Sequence 5963, Ap
25	94	3.6	387	10	US-09-789-561-156	Sequence 156, App
26	93.5	3.6	806	12	US-10-003-405-2	Sequence 2, Appl1
27	92.5	3.6	640	10	US-09-738-626-5869	Sequence 5869, Ap
28	92.5	3.6	948	10	US-09-918-951-4	Sequence 4, Appl1
29	92.5	3.6	640	10	US-10-267-311-21	Sequence 21, Appl
30	92	3.6	385	10	US-09-853-918-37	Sequence 37, Appl
31	92	3.6	532	9	US-09-966-614-2	Sequence 2, Appl1
32	91.5	3.5	1176	9	US-09-918-508-2	Sequence 2, Appl1
33	91	3.5	679	10	US-09-815-242-5407	Sequence 5407, Ap
34	91	3.5	714	9	US-10-098-807-2	Sequence 2, Appl1
35	91	3.5	730	10	US-09-815-242-112489	Sequence 12489, A
36	91	3.5	1386	10	US-09-866-582-38	Sequence 38, Appl
37	90.5	3.5	351	9	US-10-078-770-138	Sequence 158, App
38	90.5	3.5	1165	10	US-09-815-242-10744	Sequence 10744, A
39	90.5	3.5	7968	9	US-10-077-130-5	Sequence 5, Appl1
40	89.5	3.5	794	10	US-09-815-242-11829	Sequence 11829, A
41	89	3.4	1173	9	US-10-135-322-19	Sequence 19, Appl
42	89	3.4	3313	10	US-09-737-149-29	Sequence 29, Appl
43	88.5	3.4	351	10	US-09-897-214-8	Sequence 8, Appl1
44	88.5	3.4	2344	10	US-09-815-242-12713	Sequence 12713, A
45	88	3.4	633	10	US-09-815-242-13318	Sequence 13318, A

ALIGNMENTS

RESULT 1
US-09-988-863a-2
; Sequence 2, Application US/09988863a
; Patent No. US20020123427A1
; GENERAL INFORMATION:
; APPLICANT: Bayer AG
; TITLE OF INVENTION: Plant phosphomevalonate kinases
; FILE REFERENCE: la A 35 018
; CURRENT APPLICATION NUMBER: US/09/988, 863a
; CURRENT FILING DATE: 2001-11-21
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 505
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-09-988-863a-2

Query Match 100.0%; Score 2581; DB 10; Length 505;
Best local similarity 100.0%; Pred. No. 9.5e-217;
Matches 505; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MAVVASAPGKVLMTGTYLLEKPNAGLVSTNARFAIYVINEEYKPEESMANKWDYKL	60
DB	1	MAVVASAPGKVLMTGTYLLEKPNAGLVSTNARFAIYVINEEYKPEESMANKWDYKL	60
QY	61	TSPOLRESMYKSLNLTLOSASASDSRNPFEHAIQVIAAHLATERDKESLHKLL	120
DB	61	TSPOLRESMYKSLNLTLOSASASDSRNPFEHAIQVIAAHLATERDKESLHKLL	120
QY	121	QGLDITLGSNDPYSYRNQIESAGLPTLPESLTLPFAFISTFNAESNGANSKPEYAKT	180
DB	121	QGLDITLGSNDPYSYRNQIESAGLPTLPESLTLPFAFISTFNAESNGANSKPEYAKT	180
QY	181	GLSSAAMTAVVAALHLYGVVDLSDPCKEGRGSDLDVIMIAQTSICLAQGVGSG	240
DB	181	GLSSAAMTAVVAALHLYGVVDLSDPCKEGRGSDLDVIMIAQTSICLAQGVGSG	240
QY	241	FDVSCAVYGSORVYRSPVSLFAOVAVTGLPLNEVIGTILKGKMDKRTFESLPLMNL	300
DB	241	FDVSCAVYGSORVYRSPVSLFAOVAVTGLPLNEVIGTILKGKMDKRTFESLPLMNL	300

OY	301	FLGPGSGGSSTPBMWGAVKKQMSDEPKARENNONISDANLELETPLNDLSKIADHMD	360
Db	301	FLGPGSGGSSTPBMWGAVKKQMSDEPKARENNONISDANLELETPLNDLSKIADHMD	360
OY	361	VYLRYIKSCSVLTSEKKWLHAATEDINEAIIKELLEAREAMLRIIRILMRONGEAASVIEP	420
Db	361	VYLRYIKSCSVLTSEKKWLHAATEDINEAIIKELLEAREAMLRIIRILMRONGEAASVIEP	420
OY	421	ESQTQLDDSYMSAGCVLLIAGYPGAGGGEDAIFAITLGGSGTKTLQAWSSHHVALLYREDP	480
Db	421	ESQTQLDDSYMSAGCVLLIAGYPGAGGGEDAIFAITLGGSGTKTLQAWSSHHVALLYREDP	480
OY	481	HGVCLGESDPRTTCITSGVSSIHLE	505
Db	481	HGVCLGESDPRTTCITSGVSSIHLE	505

RESULT 2
US-09-909-745-22

1 Patent 22, Application US/05909745
 2 Sequence No. US20020119546A1
 3 GENERAL INFORMATION:
 4 APPLICANT: Falco, S. Carl
 5 TITLE OF INVENTION: Squalene Synthesis En-
 6 FILE REFERENCE: Bb1112 US CIP
 7 CURRENT FILING DATE: US/09/909,745
 8 PRIOR APPLICATION NUMBER: 60/07-20
 9 PRIOR FILING DATE: 1998-11-05
 10 PRIOR APPLICATION NUMBER: 60/107,241
 11 PRIOR FILING DATE: 1998-11-05
 12 NUMBER OF SEQ ID NOS: 26
 13 SOFTWARE: Microsoft Office 97
 14 SEQ ID NO: 22
 15 LENGTH: 179
 16 TYPE: PRT
 17 ORGANISM: Oryza sativa
 18 US-09-909-745-22

	Query Match	22.3%	Score 575;	DB 10;	length 179;
	Best Local Similarity	62.0%;	Pred. No. 1,8e-42;		
	Matches 111;	Conservative 26;	Mismatches 42;	Indels 0;	Gaps 0.
Oy	308 GGSPPSWAGVAKKQWMSDPEKARNMONLSDANIELETKLNDLSKLADHDVYLRVIK	367			
Dd	1 GGSSPSPWSGVSKOMQSDPOKSKETWKSGLIANSYLENQNLKNLNKLADHNEAVESYLK	60			
Oy	368 SCVSUTSEKWVLHAPEIPINEALIKELLEAREMLRIRILROMGEAASVPVEPSOTOLL	427			
Dd	61 SCSRLTCSKMTVEVANAHOELIVRSILLARDAFLETRILRMREMGIAAGVIPESPOTOLL	120			
Oy	428 DSTMAEGVLLAGVGAGGFDPAIFATIGDSCTKTLQAAMSHNYALLVRDEPHCYCE	486			
Dd	121 DATMMEGVLLAGVGAGGFDPAVFVVYLLGAASDAVAKAASAGVPLVRDEPRGVSLK	179			

RESULT 3
 US-09-909-745-24
 ; Sequence 24, Application US/09909745
 ; Patent No. US20020119546v1
 ; GENERAL INFORMATION:
 ; APPLICANT: Falco, S. Carl
 ; APPLICANT: Farnodu, Omolayo O.
 ; TITLE OF INVENTION: Squelene Synthesis Enzymes
 ; FILE REFERENCE: BB112 US CIP
 ; CURRENT APPLICATION NUMBER: US/09/909,745
 ; CURRENT FILING DATE: 2001-07-20
 ; PRIOR APPLICATION NUMBER: 60/107,241
 ; PRIOR FILING DATE: 1998-11-05
 ; PRIOR APPLICATION NUMBER: 60/107,241
 ; PRIOR FILING DATE: 1998-11-05
 ; NUMBER OF SEQ ID NOS: 26

```

; SOFTWARE: Microsoft Office 97
; SEQ ID NO 24
; LENGTH: 451
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-09-909-745-24

```

Query Match	17.6%;	Score 453.5;	DB 10;	Length 451;
Best Local Similarity	30.1%;	Pred. No. 2.9e-31;		
Matches 155;	Conservative 72;	Mismatches 185;	Indels 103;	Gaps 18

```

0Y      6 SAPGVMTGGVLYLEKRNAGLVSTANRFAAIVKPINE-----EVKEESWAMK--W 55
          ||||| : ||||| : : : | : : | : : | : : |
Db      8 SAPGALLAGGYLVDFTEAFVVGGLSARMAHAHPYGSLGSDXKFVRVRSKQFKGEW 67
          ||||| : ||||| : : : | : : | : : | : : |
0Y      56 TDVKLTSPOLSRBSNKYSILNHLTLOS---VSASDSRNPPEVAIOYAIAAAHLLATEKD 111
          || : : : ||||| : : : | : : | : : | : : |
Db      68 -----LTHISPKSGGFIPVSGIGSKNPFIEKVIANVFSYF----- 101
0Y      112 KESLHKLLLOGLDITIIIGSNFYGNONIESAGLPRLPESLGTLAPASTFPMAAESNGA 171
          || : : : ||||| : : | : : | : : | : : | : : |
Db      102 KPNNDDCYNRULEVIDIFSDV--AIHSOEDS-----VTEHRG-----NRRLSFHS 144
0Y      172 NSKPEVATGTGSSSAAMTTAVAAALHNLYGLVDLDSECKEKGFCSDLDVIHMIAQTSHC 231
          ||||| : ||||| : : : | : : | : : | : : | : : |
Db      145 HRIEEVPRTGTGSAGLVTVITTLALSPF--VSDLENNVDKR-----EVIHNLAQVAHC 197
0Y      232 LAQGCVSGEPVSCAVYGSGORRYRPSREVLSPAQVAVTGLPLNEVITTIKKK----- 284
          ||||| : ||||| : : : | : : | : : | : : | : : |
Db      198 QAOGKIGSGEPVAAAAYGISIRYREFPALIS-----NLP--DISARYGSKIHLVD 247
0Y      285 ---MDNKRTETSLPMLMLFLGPEGSGSGSSPPSMVGAVKKMKMSDEPKARENNOJSDAN 341
          || : : : ||||| : : : | : : | : : | : : | : : |
Db      248 EEDMNTTKTSNHLPSGLTLMWGD-I-KNSGETVKLVQYKKNKYDSHMPESLKITYELDHAN 306
0Y      342 LELEFKINDLSKLAKDHMDVLYARIKSC--SVLTSEKVMYLHATERINEALIKELLEAREA 399
          || : : : ||||| : : | : : | : : | : : | : : |
Db      307 SRFMDOGLSKDLRHETHHDYSDQJFESLERNDCTCOXY-----PETTEVEDA 353
0Y      400 MLRFILMRONGEAASVIPESPOTQLDDSTMSAEVLYLAGVPAGGFDIAIATLTDSC 455
          || : : : ||||| : : | : : | : : | : : | : : |
Db      354 VATIRRSRKTYTKESGADIEPPVOTSLLDCQTLKGVLTCLIPAGGYDAIAVITKODVD 413
0Y      460 TKLTQWASH--NVYALLYREDPHGCLESGEDPT 492
          || : : : ||||| : : | : : | : : | : : | : : |
Db      414 LRAQTANDKRFASKYOIMLDVTOADNGVKKRK--DPELT 447

```

```

RESULT
US-09-909-745-20
; Sequence 20, Application US/09909745
; Patent No. US20020119546A1
; GENERAL INFORMATION:
; APPLICANT: Falco, S. Carl
; TITLE OF INVENTION: Squalene Synthesis Enzymes
; FILE REFERENCE: BB112 US CIP
; CURRENT APPLICATION NUMBER: US/09/909, 745
; PRIOR FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: 60/107, 241
; PRIOR FILING DATE: 1998-11-05
; PRIOR APPLICATION NUMBER: 60/107, 241
; PRIOR FILING DATE: 1998-11-05
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: Microsoft Office 97
; SEQ ID NO: 20
; LENGTH: 67
; TYPE: prt
; ORGANISM: Oryza sativa
US-09-909-745-20

```

Query Match	8.48;	Score 216;	DB 10;	Length 677
Best Local Similarity	60.08;	Pred. No. 8.5e-12;		

Matches 39; Conservative 12; Mismatches 14; Indels 0; Gaps 0;

OY 309 GSSPPSWGAVKRWOMSPDEKARWMLSDANLELETKLNDLSKIAKHHDVYLRTVKS 368
Db 1 GSSPPSWGAVKRWOMSPDEKARWMLSDANLELETKLNDLSKIAKHHDVYLRTVKS 60

OY 369 CSVLT 373
Db 61 CSRLT 65

RESULT 5
US-09-909-745-16
; Sequence 16, Application US/09909745
; Patent No. US20020119546A1
; GENERAL INFORMATION:
; APPLICANT: Falco, S. Carl
; APPLICANT: Famodu, Omolayo O.
; TITLE OF INVENTION: Squalene Synthesis Enzymes
; FILE REFERENCE: BB1112 US CIP
; CURRENT APPLICATION NUMBER: US/09/909,745
; CURRENT FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: 60/107,241
; PRIOR FILING DATE: 1998-11-05
; PRIOR APPLICATION NUMBER: 60/107,241
; PRIOR FILING DATE: 1998-11-05
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 16
; LENGTH: 54
; TYPE: PRT
; ORGANISM: Zea mays
US-09-909-745-16

Query Match
Best Local Similarity 74.1%; Pred. No. 6.8e-11;
Matches 40; Conservative 8; Mismatches 6; Indels 0; Gaps 0;

OY 397 REAMLRIRILRMKGAEASVPIEPESOTQLDSTMSAGVLLAGVPGAGGPDAT 450
Db 1 RDACLEIRLHREMGIAAGVPIEPESOTRLDATTMNGVLLAGVPGAGGPDAT 54

RESULT 6
US-09-909-745-18
; Sequence 18, Application US/09909745
; Patent No. US20020119546A1
; GENERAL INFORMATION:
; APPLICANT: Falco, S. Carl
; APPLICANT: Famodu, Omolayo O.
; TITLE OF INVENTION: Squalene Synthesis Enzymes
; FILE REFERENCE: BB1112 US CIP
; CURRENT APPLICATION NUMBER: US/09/909,745
; CURRENT FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: 60/107,241
; PRIOR FILING DATE: 1998-11-05
; PRIOR APPLICATION NUMBER: 60/107,241
; PRIOR FILING DATE: 1998-11-05
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 18
; LENGTH: 54
; TYPE: PRT
; ORGANISM: Zea mays
US-09-909-745-18

Query Match
Best Local Similarity 74.1%; Pred. No. 6.8e-11;
Matches 40; Conservative 8; Mismatches 6; Indels 0; Gaps 0;

OY 397 REAMLRIRILRMKGAEASVPIEPESOTQLDSTMSAGVLLAGVPGAGGPDAT 450
Db 1 RDACLEIRLHREMGIAAGVPIEPESOTRLDATTMNGVLLAGVPGAGGPDAT 54

RESULT 7
US-09-815-242-10656
; Sequence 10656, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10656
; LENGTH: 368
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-815-242-10656

Query Match
Best Local Similarity 22.3%; Pred. No. 2.1e-07;
Matches 106; Conservative 58; Mismatches 179; Indels 133; Gaps 20;

OY 3 VVASAPGVMTGTYLEKRPAGLVSTNARFAIVPIEVEVPEPMARKMTIVKLS 62
Db 2 IEVTPGKFLIAGEYAVPEGHPALIVADQFVTVETDEGSIOGAOYSSLPFRWT 60
OY 63 POLRESMYKLSLNLTLQSVASDSRNPVEHAIOVAIAAHATEKDKESLHKLLOG 122
Db 61 ---RNGELVDIR-----ENPF-----HYVLAHL-TEKTAOQNKEL--- 96
OY 123 LDITLLGNDPFSYRNOIESAGLPLTPESLGLAPFASITFNAASNGANSKPEVAKTGL 182
Db 97 -----SFYHLK-----VTSLSDSNG-----RKYGL 117
OY 183 GSSAMTTAAVAAL--LHYLGAVDSDPCKEKCEGSDLDVTHMAQTSKLAQKVS 240
Db 118 GSSGAVTVGTAKALNIFDGL-----ENEFKLSALAHAVOQN-GSC 161
OY 241 FDVSCAVYGSGRYRSPSEVLISFAOVAATGLPLNEVIGITLKGDKNKRT-EFSLPPLMN 299
Db 162 GDIAASCTGG--WIAFTFDHDMVQKYT-----TETLDLAMPPELMIPPLKVPKQIR 215
OY 300 LFLGPGGSGSTPSMGAVKRWOMSPDEKARWMLSDANLELETKLNDLSKIAKHHDV 359
Db 216 LLIGWTGS-PASTSLDVRV---HOSKEKQAAVEQPLMKSLCYETMINGN----- 264
OY 360 DYTLAVITSCSVLISEKWLHATEPINEAIIKELLEAREAMLRIRILRMKGAEASVPI 419
Db 265 -----TGRIS-VIOKQITKRNQ-----LTAETSLTGAVIE 294

OY 420 PESOTQLDSTMSABEYLLAGVPGAGFPAIFAIITLGDG-TKITQAMSHNVLA 474
DB 295 TEALKNLCIDIAESYTG--AKSSGAGGDCGIYIFROKSGILPLMTAMEGDITPL 348

RESULT 8

US-09-815-242-12866
Sequence 12866, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlssen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
PRIOR FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for windows Version 4.0
SEQ ID NO 12866
LENGTH: 358
TYPE: PRF
ORGANISM: Staphylococcus aureus
US-09-815-242-12866

Query Match 5.2%; Score 135.5; DB 10; Length 358;
Best Local Similarity 22.6%; Pred. No. 0.0011;
Matches 108; Conservative 58; Mismatches 178; Indels 133; Gaps 25;

OY 3 VVASAPGKVLMTGGVLYLEKPNAGVLSTNARFYAIKPIINEEYKPPESMAKMTDVKLT 62
DB 2 IQVAKGCKLTIAGEYAVTE--PGYKSVYLALDR--VYATIEE-----TD----- 42
OY 63 POLSRESMYLSLNHLTLOSASDSRNPVEHAIOYAIANAHLATEKDESLHKLLOG 122
DB 43 -----QIKGTHSKALHNPVTFSRD--EDSI--VISDPHAKQ-----TD----- 77
OY 123 LDTITLGSNDYFSRNOIESA--GLPLTPESLGLTAPFASITFNAESNGANSKPEYAKT 180
DB 78 LNVVVAIEIFEQYAKSCDIAMKHFHLLTDS-----NLDSNG-----HKY 118
OY 181 GLGSSAAMTTAAVVAALHLYGVVDLSDPCKEKGCSDDLVIHMLAQTSHCLAQKVGSG 240
DB 119 GLGSSAAVLVSIVIKVINEFDM-----KLSNL--YIKKLVIANMKLOSLSG 165
OY 241 FDSVCAY--GSQRYVRPSPVLSFAQAVYGLPLNEVIGTILKGKMDNKRTF--FSLPPLM 298
DB 166 -DIAVSVISGMALSTFDHEWVKH-QIEDT-----TVEEVLKINMGGLHIEPLQAPENM 217
OY 299 NLELPGSGSGSTPMSVAGVKKWQMSDPEKARENMONLSDANILETKLNDLSKLAKDH 358
DB 218 EVLIGMTGSPASS--PHFVSEVKRLK--SDP-----SF 246

OY 359 WDYYLRVYISCSVLTSEKVLHATEPINEAIIELLEAREAMRIRILMROGEAASVPI 418
DB 247 YCDFLEDSHRC-----VEK-LIHAFKTNINIKGVQKVRON-----RTIIORMKREATVDI 295
OY 419 EPESQQLDSTMSABEYLLAGVPGAGFPAIFAIITLGD--SGTKLQAMSHNVLA 474
DB 296 ETEKRLKYLCDIAEKYHGA--SKTSGAGGDCGITTITNKVDYDKERTYDEMTKHGIKPL 350

RESULT 9

US-09-815-242-5854
Sequence 5854, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlssen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
PRIOR FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for windows Version 4.0
SEQ ID NO 5854
LENGTH: 345
TYPE: PRF
ORGANISM: Staphylococcus aureus
US-09-815-242-5854

Query Match 5.2%; Score 133.5; DB 10; Length 345;
Best Local Similarity 22.8%; Pred. No. 0.0016;
Matches 107; Conservative 57; Mismatches 173; Indels 133; Gaps 25;

OY 7 APKRVLMTGGVLYLEKPNAGVLSTNARFYAIKPIINEEYKPPESMAKMTDVKLTSPOLS 66
DB 3 APKRVLMTGGVLYLEKPNAGVLSTNARFYAIKPIINEEYKPPESMAKMTDVKLTSPOLS 66
OY 67 RESMYKLSLNHLTLOSASDSRNPVEHAIOYAIANAHLATEKDESLHKLLOGLDIT 126
DB 40 ---QYKGTIHSKALHNPVTFSRD--EDSI--VISDPHAKQ-----LNVY 78
OY 127 ILGSNDYFSRNOIESA--GLPLTPESLGLTAPFASITFNAESNGANSKPEYAKTGLG 184
DB 79 VTAIEIFEQYAKSCDIAMKHFHLLTDS-----NLDSNG-----HKYGLG 119
OY 185 SAAMTTAAVVAALHLYGVVDLSDPCKEKGCSDDLVIHMLAQTSHCLAQKVGSGFDS 244
DB 120 SAADVLSIVIKVINEFDM-----KLSNL--YIKKLVIANMKLOSLSG--DIA 165
OY 245 CAVY--GSQRYVRPSPVLSFAQAVYGLPLNEVIGTILKGKMDNKRTF--FSLPPLANLFL 302
DB 166 VSVISGMALSTFDHEWVKH-QIEDT-----TVEEVLKINMGGLHIEPLQAPENNEVL 218

```

OY 303 GEPGGGGSTPMGYAKKQNMSPBEKAREMOWULSDANIELEKRLKNDLSKLKDHNDY 366
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 219 GWTSPASS-PHFPSEYKRLK-SDP-----SFYGD 247
OY 363 LRIVKSCVLSSEKVVHLATEPIEATITIKELLEAREAMLRIILROMGEASVPIEPES 422
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 248 LEDSHRC---VER-LTHAFETNNIKGVQKWVRN-----RTIIQRDRKEATVDIETEK 296
OY 423 QOTLLDSTMSAEGVLAGVPAGEFDAIFATTLGD-SGTQLTQWSSSHNV 471
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 297 LKYLCDIAEKYHGA--SKTSGAGGDCGITYINKDVREKIEYDEWTHNGI 344

```

RESULT 10
 US-09-909-745-26
 : Sequence 26, Application US/09909745
 : Patent No. US20020119546A1
 : GENERAL INFORMATION:
 : APPLICANT: Falco, S. Carl
 : APPLICANT: Famodu, Omolayo O.
 : TITLE OF INVENTION: Squalene Synthesis Enzymes
 : FILE REFERENCE: BB112 US CIP
 : CURRENT APPLICATION NUMBER: US/09/909,745
 : CURRENT FILING DATE: 2001-07-20
 : PRIOR APPLICATION NUMBER: 60/107,241
 : PRIOR FILING DATE: 1998-11-05
 : PRIOR APPLICATION NUMBER: 60/107,241
 : PRIOR FILING DATE: 1998-11-05
 : NUMBER OF SEQ ID NOS: 26
 : SOFTWARE: Microsoft Office 97
 : SEQ ID NO 26
 : LENGTH: 386
 : TYPE: PRT
 : ORGANISM: Glycine max
 : US-09-909-745-26

Query Match	4.9%	Score 126;	DB 10;	Length 386;
Best Local Similarity	21.0%;	Pred. No. 0.0084;		
Matches 97;	Conservative 73;	Mismatches 155;	Indels 136;	Gaps

[illegible]

```

RESULT 11
US-09-909-745-23
; Sequence 23, Application US/09909745
; Patent No. US20020119546A1
; GENERAL INFORMATION:
; APPLICANT: Falco, S. Carl
; TITLE OF INVENTION: Squalene Synthesis Enzymes
; FILE REFERENCE: BR112 US CIP
; CURRENT APPLICATION NUMBER: US/09/909,745
; CURRENT FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: 60/107,241
; PRIOR FILING DATE: 1998-11-05
; PRIOR APPLICATION NUMBER: 60/107,241
; PRIOR FILING DATE: 1998-11-05
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 23
; LENGTH: 378
; TYPE: PRF
; ORGANISM: Arabidopsis thaliana
US-09-909-745-23

```

Query Match	4.88;	Score 123;	DB 10;	Length 378;
Best Local Similarity	20.98;	Pred. No. 0.015;		
Matches 100; Conservative	78;	Mismatches 151;	Indels 150;	Gaps 23;

[illegible]

RESULT 12
US-09-925-388-8
Sequence 8, Application US/09925388
Publication No. US20030054523A1
GENERAL INFORMATION:
APPLICANT: HOSHINO, Tatsuo
APPLICANT: OJIMA, Kazuyuki
APPLICANT: SETOGUCHI, Yutaka
TITLE OF INVENTION: ISOPRENOID PRODUCTION
FILE REFERENCE: ISOPRENOID PRODUCTION
CURRENT APPLICATION NUMBER: US/09/925,388
CURRENT FILING DATE: 2001-08-09

;; PRIOR APPLICATION NUMBER: 09/306,595
 ;; PRIOR FILING DATE: 1999-05-06
 ;; NUMBER OF SEQ ID NOS: 43
 ;; SOFTWARE: PatentIn Ver. 2.1
 ;; SEQ ID NO 8
 ;; LENGTH: 432
 ;; TYPE: PRT
 ;; ORGANISM: Phaffia rhodozyma
 ;; US-09-925-388-8

Query Match 4.5%; Score 115; DB 9; Length 432;
 Best Local Similarity 21.6%; Pred. No. 0.091; Indels 120; Gaps 19;
 Matches 103; Conservative 64; Mismatches 190;

QY 3 VVASAPGVMTGCVLYLEKPNAGLVLTNARFAIYKPIINEEVKPESMAMKMTDVYKLS 62
 DB 4 ILVSAPGVKILFGEHVA-GHGVYGIASVPLRCALLSP-----TATTTTS 48
 QY 63 POLSRESMYKLSLNHLTLOSASDSRNPEVAIOYAIAMHLATEKD-KESLHKLLQ 121
 DB 49 SLSLSYNI-TISLTLNFTQSWPVS-----LPMSLAPDW--TEASIPESICPTLLA 97
 QY 122 GLD-TITLGSND-----FYSYRQIESAGLPLTPESIGTLAPASTFMAAESNGA 171
 DB 98 EIERIAGGGNGGEREVATMAFLYLVLVLSKGRPEPEL-----TANSALPMGA 148
 QY 172 NSRPEVAKTGLSSAAMTTAVALLHYLGVDLSDPCKEKGFGCSDDLYIHIAQTSHC 231
 DB 149 -----GLSSAALSTSLALVFLFHSLSPTTGRESTIPTADTEVIDKWAFLAEK 199
 QY 232 LAGKVGSGFDVSCAVYGSRYVRFSEPVLSFAVANTGLPLNEVICTILKGKDKNRT 291
 DB 200 VTHGN-PSGIDNNAVSTRGG-----AAAEKRIEKGQ-E 230
 QY 292 FSLPPLMANL---FLGEPGSGSSTPSMGAVKMKQMSDEK-----ARENWON 336
 DB 231 GMEAIKSFSTIRFLTDSIGRDTSLVAGVANRLQEPVIVPLEAQIQADEAIRC 290
 QY 337 LSDANLETKLNDLSKLAKDHDVYLRIKSCSVLTSEKWLHATEPIEAIKELLEA 396
 DB 291 LKDSERAMVIDRLQNLVSEN-----HAHLAALGVSHPSLEIIR-IGAD 335
 QY 397 REMLRIRILRMGGEASVPIEPESOTOLLDSIM-----SABEVLLAGPGAG-GF 447
 DB 336 KPPELRKLTGAGGCGCAVTLVPDDSTETLQALMETLVGSSPAPYIARVGGSGVGF 392

RESULT 13

;; US-10-092-880-4
 ;; Sequence 4, Application US/10092880
 ;; Patent No. US20020164354A1
 ;; GENERAL INFORMATION:
 ;; APPLICANT: Barenkamp, Stephen J.
 ;; TITLE OF INVENTION: HIGH MOLECULAR WEIGHT SURFACE PROTEINS OF NON-TYPEABLE
 ;; FILE REFERENCE: HAEMOPHILUS
 ;; CURRENT APPLICATION NUMBER: US/10/092,880
 ;; CURRENT FILING DATE: 2002-03-08
 ;; PRIOR APPLICATION NUMBER: 09/155,614
 ;; PRIOR FILING DATE: 1998-09-30
 ;; PRIOR APPLICATION NUMBER: 08/617,697
 ;; PRIOR FILING DATE: 1996-04-01
 ;; PRIOR APPLICATION NUMBER: PCT/US97/04707
 ;; PRIOR FILING DATE: 1997-04-01
 ;; NUMBER OF SEQ ID NOS: 11
 ;; SOFTWARE: PatentIn Ver. 2.1
 ;; SEQ ID NO 4
 ;; LENGTH: 1477
 ;; TYPE: PRT
 ;; ORGANISM: Haemophilus influenzae
 ;; US-10-092-880-4

Query Match 4.2%; Score 108.5; DB 9; Length 1477;

Best Local Similarity 20.3%; Pred. No. 2.2;
 Matches 89; Conservative 68; Mismatches 162; Indels 119; Gaps 21;

QY 11 VLMTGGVLYLEKPN-----GLVLTNARFAIYKPIINEEVKPESMAMKMT----- 56
 DB 139 ILDSNGCVFLINPNCITIGDAIINTNGFTASTLUDISNENIKARNFTFEDTKKALAEIV 198
 QY 57 -----DVKLTSPOLSRSMYKL---SLNHLTLOSASDSRNPEVAIOYAI 102
 DB 199 NHGLITVKGDSVNLIGKGVKNEGVIVSNGSISLLAGKITISDIINP-----TIFYSIA 254
 QY 103 AAILATEKDESLHKLLOGDITLIGSNDIFYRQIESAGLPLPESIGTLAPASTIT 162
 DB 255 -----APENEAVNLGDIFFAKGNINRAA---TIRQNGKLSADSVSKDKGNIV--LSAK 304
 QY 163 FNAESNGA-NSRPEVAKT-----LGSSAAMTTAVAAALL-----HYLGVDLSPCKE 211
 DB 305 EGEAEIIGVISAQNOQAKGKMLITGDKYTLKIGAVIDLSGKGGETYLQ-----GGERGE 360
 QY 212 GKEGCSDDLYIHIAQT-----SHCLAQKVGSGFDVSCAVYGSRYVRFSEPVLSFAOY 266
 DB 361 KRNG-----IQLAKRTSLEKSGTINVSKEKGFAL---VMGDIALIDGNINAQSGDI 411
 QY 267 AVTG-----LPLNEVICTILKGK-----DNKRTFESLPLMNLFLGEPGSGGS 311
 DB 412 AKTGGEVETSGHDLFTKDNALVDA---KEWLLDPDNVSIABDPFNNGINDPEPTGTG 468
 QY 312 TPSMGAVKMKQMSDEKARENQNLSDANLETKLND--LSKLAKDHDV-----YL 363
 DB 469 -----EASDPK-----NSELKTTLTNTTISNYLKNAMTMTITASRKL 506
 QY 364 RVIKSCSVLTSEKWLHA 381
 DB 507 TVNSINIGSNHLLILHS 524

RESULT 14

;; US-10-077-023-7
 ;; Sequence 7, Application US/10077023
 ;; Publication No. US20030031675A1
 ;; GENERAL INFORMATION:
 ;; APPLICANT: MIKESSELL, GLEN E.
 ;; APPLICANT: CHANG, HAN
 ;; APPLICANT: FINGER, JOSHUA N.
 ;; APPLICANT: YANG, GUOCHEN
 ;; APPLICANT: LU, PIN
 ;; APPLICANT: ZHOU, XIA-DI
 ;; APPLICANT: PEACH, ROBERT
 ;; TITLE OF INVENTION: B7-RELATED NUCLEIC ACIDS AND POLYPEPTIDES USEFUL FOR
 ;; FILE REFERENCE: IMMUNOMODULATION
 ;; CURRENT APPLICATION NUMBER: US/10/077,023
 ;; CURRENT FILING DATE: 2002-02-15
 ;; PRIOR APPLICATION NUMBER: 60/272,107
 ;; PRIOR FILING DATE: 2001-02-28
 ;; PRIOR APPLICATION NUMBER: 60/209,811
 ;; PRIOR FILING DATE: 2000-06-06
 ;; NUMBER OF SEQ ID NOS: 138
 ;; SOFTWARE: PatentIn Ver. 2.1
 ;; SEQ ID NO 7
 ;; LENGTH: 534
 ;; TYPE: PRT
 ;; ORGANISM: Homo sapiens
 ;; US-10-077-023-7

Query Match 3.9%; Score 101.5; DB 9; Length 534;

Best Local Similarity 22.8%; Pred. No. 1.9; Indels 145; Gaps 23;
 Matches 97; Conservative 48; Mismatches 136;

QY 183 GSSAAMTTAVAAALHY-----LGAVDLSDPCKEKGFGC-----SDLDVYIHM 225
 DB 86 GSAVANRTALFPDLAAGNASLRLQVRVAD---ESFICFVSIRDFGSAAVSLQVAAPY 142

QY 226 AOTSCHLAOCK---VSGGFVSCAVYGSQRYVRFSEVLSPFAQVAVTGLPLNEVIGTILK 282
DB 143 SKPSMTLEPKDLRPGDTVTITCSSY--QGY-----PEAEVFMQDG--QGVPLT---GNVTT 192
QY 283 GKDNKTFEESLPLPLNLFGEPC-----SGGSSNP--SMGCAVKKWQM 324
DB 193 SOMANEGLDFVHSILRVILGANGTYSCLVRNPVLQODAHSSVTITPQRSPTGAVEQVP 252
QY 325 SDPEKARENQONSDANILEETKLNLSKLAKHMDVYLIRVYKSCSVLTSEKWLHA-TE 383
DB 253 EDPVVALVG---TDATLRCSFSPEPGFSILAQ-----LNLIMO---LITDKLVHSFTE 299
QY 384 PINE-----AIKELLEAREAMLIR-----ILROMGEAA----- 414
DB 300 GRDGSAYANRTALFPDLLAOGNASLRLQVRVADEGSFTCFVSIIRFGSAVSLQVAP 359
QY 415 ----SVPIEESQTQLDST-----MSAEVLLAG-----VPAGG 446
DB 360 YSKPSMTLEPKDLRPGDTVTITCSSYRGYPEAEVFMQDGGVPLTGNVTTSQMANEGL 419
QY 447 FD--AIFAITLGDGSKTLTQAMSSHNVLALLVR-----EDPHGVCLESGDPRT-----TC 494
DB 420 EDVHSVLRVILGANGT-----YSCLVARNPVLQODAHSSVTITGQPMTPPPEALW 468
QY 495 ITSGVS 500
DB 469 VTVGLS 474

RESULT 15
US-09-875-338-7
; Sequence 7, Application US/09875338
; Patent No. US20020095024A1
; GENERAL INFORMATION:
; APPLICANT: MISELL, GLEN E.
; APPLICANT: CHANG, HAN
; APPLICANT: FINGER, JOSHUA N.
; APPLICANT: YANG, GUCHEN
; APPLICANT: LU, PIN
; APPLICANT: ZHOU, XIA-DI
; APPLICANT: PEACH, ROBERT
; TITLE OF INVENTION: B7-RELATED NUCLEIC ACIDS AND POLYPEPTIDES USEFUL FOR
; FILE REFERENCE: 3053-4071US2
; CURRENT APPLICATION NUMBER: US/09/875,338
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: 60/272,107
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/209,811
; PRIOR FILING DATE: 2000-06-06
; NUMBER OF SEQ ID NOS: 94
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 534
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-875-338-7

Query Match 3.9%; Score 101.5; DB 10; Length 534;
Best Local Similarity 22.8%; Pred. No. 1.9;
Matches 97; Conservative 48; Mismatches 136; Indels 145; Gaps 23;

QY 183 GSSAAMTTAVVALLHR-----LGVVDLSDPCKEKGFGC-----SDLDVIHMI 225
DB 86 GSAVANRTALFPDLLAOGNASLRLQVRVAD---EGSFTCFVSIIRFGSAVSLQVAPY 142
QY 226 AOTSCHLAOCK---VSGGFVSCAVYGSQRYVRFSEVLSPFAQVAVTGLPLNEVIGTILK 282
DB 143 SKPSMTLEPKDLRPGDTVTITCSSY--QGY-----PEAEVFMQDG--QGVPLT---GNVTT 192
QY 283 GKDNKTFEESLPLPLNLFGEPC-----SGGSSNP--SMGCAVKKWQM 324
DB 193 SOMANEGLDFVHSILRVILGANGTYSCLVRNPVLQODAHSSVTITPQRSPTGAVEQVP 252

QY 325 SDPEKARENQONSDANILEETKLNLSKLAKHMDVYLIRVYKSCSVLTSEKWLHA-TE 383
DB 253 EDPVVALVG---TDATLRCSFSPEPGFSILAQ-----LNLIMO---LITDKLVHSFTE 299
QY 384 PINE-----AIKELLEAREAMLIR-----ILROMGEAA----- 414
DB 300 GRDGSAYANRTALFPDLLAOGNASLRLQVRVADEGSFTCFVSIIRFGSAVSLQVAP 359
QY 415 ----SVPIEESQTQLDST-----MSAEVLLAG-----VPAGG 446
DB 360 YSKPSMTLEPKDLRPGDTVTITCSSYRGYPEAEVFMQDGGVPLTGNVTTSQMANEGL 419
QY 447 FD--AIFAITLGDGSKTLTQAMSSHNVLALLVR-----EDPHGVCLESGDPRT-----TC 494
DB 420 EDVHSVLRVILGANGT-----YSCLVARNPVLQODAHSSVTITGQPMTPPPEALW 468
QY 495 ITSGVS 500
DB 469 VTVGLS 474

Search completed: April 26, 2003, 13:02:22
Job time : 22 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4_p5.4578
Copyright (c) 1993 - 2003 Comphen Ltd.

OM protein - protein search, using sw model

Run on: April 26, 2003, 12:55:10 ; Search time 147 seconds
(without alignments)
2214.899 Million cell updates/sec

Title: US-09-988-863a-2
Perfect score: 2581
Sequence: 1 MAVVASAPGKVLMTGCVL.....ESGDPRTCTITGSVSHLE_505

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 4569144 seqs, 644733110 residues

Total number of hits satisfying chosen parameters: 4569144

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Pending_Patents_AA_Main:*
1: /cgn2_6/ptodata/1/paa/PCRTUS_COMB.pep:*
2: /cgn2_6/ptodata/1/paa/US06_COMB.pep:*
3: /cgn2_6/ptodata/1/paa/US07_COMB.pep:*
4: /cgn2_6/ptodata/1/paa/US08_COMB.pep:*
5: /cgn2_6/ptodata/1/paa/US081_COMB.pep:*
6: /cgn2_6/ptodata/1/paa/US082_COMB.pep:*
7: /cgn2_6/ptodata/1/paa/US083_COMB.pep:*
8: /cgn2_6/ptodata/1/paa/US084_COMB.pep:*
9: /cgn2_6/ptodata/1/paa/US085_COMB.pep:*
10: /cgn2_6/ptodata/1/paa/US086_COMB.pep:*
11: /cgn2_6/ptodata/1/paa/US087_COMB.pep:*
12: /cgn2_6/ptodata/1/paa/US088_COMB.pep:*
13: /cgn2_6/ptodata/1/paa/US089_COMB.pep:*
14: /cgn2_6/ptodata/1/paa/US090_COMB.pep:*
15: /cgn2_6/ptodata/1/paa/US091_COMB.pep:*
16: /cgn2_6/ptodata/1/paa/US092_COMB.pep:*
17: /cgn2_6/ptodata/1/paa/US093_COMB.pep:*
18: /cgn2_6/ptodata/1/paa/US094_COMB.pep:*
19: /cgn2_6/ptodata/1/paa/US095_COMB.pep:*
20: /cgn2_6/ptodata/1/paa/US096_COMB.pep:*
21: /cgn2_6/ptodata/1/paa/US097_COMB.pep:*
22: /cgn2_6/ptodata/1/paa/US098_COMB.pep:*
23: /cgn2_6/ptodata/1/paa/US099_COMB.pep:*
24: /cgn2_6/ptodata/1/paa/US100_COMB.pep:*
25: /cgn2_6/ptodata/1/paa/US101_COMB.pep:*
26: /cgn2_6/ptodata/1/paa/US102_COMB.pep:*
27: /cgn2_6/ptodata/1/paa/US60_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2581	100.0	505	US-09-988-863a-2	Sequence 2, Appl1
2	1930.5	74.8	503	PCT-US02-24048A-12	Sequence 12, Appl
3	1930.5	74.8	503	US-10-036-959B-12	Sequence 12, Appl
4	1579	61.2	511	US-10-219-999-40847	Sequence 40847, A
5	1579	61.2	511	US-60-324-109-16899	Sequence 16899, A
6	1531	59.3	499	US-09-708-427-83408	Sequence 83408, A

	7	1531	59.3	535	21	US-09-708-427-83407	Sequence 83407, A
	8	1341	52.0	535	21	US-09-708-427-66289	Sequence 66289, A
	9	973	37.7	326	21	US-09-708-427-66290	Sequence 66290, A
	10	973	37.7	326	21	US-09-708-427-83409	Sequence 83409, A
	11	902	34.9	306	21	US-09-708-427-66291	Sequence 66291, A
	12	868.5	33.6	286	26	US-10-219-999-55015	Sequence 55015, A
	13	839.5	32.5	281	26	US-10-219-999-42346	Sequence 42346, A
	14	839.5	32.5	281	27	US-60-312-544-6223	Sequence 6223, Ap
	15	722.5	28.0	359	21	US-09-708-427-65887	Sequence 65887, A
	16	612.5	23.7	501	27	US-60-341-261-3218	Sequence 3218, Ap
	17	575	22.3	179	18	US-09-433-242-22	Sequence 22, Appl
	18	575	22.3	179	18	US-09-433-242A-22	Sequence 3740, Appl
	19	575	22.3	179	23	US-09-909-745-22	Sequence 22, Appl
	20	529	20.5	402	27	US-60-360-039-3740	Sequence 3740, Appl
	21	492.5	19.1	451	1	PCT-US02-13142-3406	Sequence 3406, Ap
	22	492.5	19.1	451	25	US-10-128-714-3406	Sequence 3406, Ap
	23	492.5	19.1	451	27	US-60-316-362-3406	Sequence 3406, Ap
	24	490	19.0	484	1	PCT-US02-13142-8406	Sequence 8406, Ap
	25	490	19.0	484	25	US-10-128-714-8406	Sequence 8406, Ap
	26	490	19.0	484	27	US-60-295-899-177	Sequence 177, App
	27	485	18.8	423	27	US-60-360-039-2162	Sequence 2162, Ap
	28	480	18.6	432	1	PCT-US02-03987-14967	Sequence 14967, A
	29	480	18.6	432	21	US-09-792-024-104	Sequence 104, App
	30	480	18.6	432	24	US-10-032-585-7023	Sequence 7023, Ap
	31	480	18.6	432	24	US-10-072-851-14967	Sequence 14967, A
	32	480	18.6	432	27	US-60-259-128-4975	Sequence 4975, Ap
	33	480	18.6	432	27	US-60-314-050-7023	Sequence 7023, Ap
	34	478	18.5	453	25	US-10-179-131-6401	Sequence 6401, Ap
	35	478	18.5	453	25	US-09-673-784A-13279	Sequence 13279, A
	36	475	18.4	432	24	US-10-069-062-7	Sequence 7, Appl1
	37	459.5	17.8	409	16	US-09-248-796-17922	Sequence 17922, A
	38	459.5	17.8	409	27	US-60-096-409-17922	Sequence 17922, A
	39	453.5	17.6	451	18	US-09-433-242-24	Sequence 24, Appl
	40	453.5	17.6	451	18	US-09-433-242A-24	Sequence 24, Appl
	41	453.5	17.6	451	23	US-09-909-745-24	Sequence 24, Appl
	42	453.5	17.6	451	27	US-60-360-039-1509	Sequence 1509, A
	43	434	16.8	153	20	US-09-617-682A-15100	Sequence 15100, A
	44	414.5	16.1	384	18	US-09-417-507-28517	Sequence 28517, A
	45	404	15.7	136	26	US-10-219-999-33417	Sequence 33417, A

ALIGNMENTS

RESULT 1
US-09-988-863a-2
; Sequence 2, Application US/0998863A
; GENERAL INFORMATION:
; APPLICANT: Bayer AG
; TITLE OF INVENTION: Plant phosphomevalonate kinases
; FILE REFERENCE: Le A 35 018
; CURRENT APPLICATION NUMBER: US/09/988, 863A
; CURRENT FILING DATE: 2001-11-21
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 505
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-09-988-863a-2

Query Match 100.0%; Score 2581; DB 23; Length 505;
Best Local Similarity 100.0%; Pred. No. 1; le-246;
Matches 505; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	DB	QY	DB
1	1	61	61
MAVVASAPGKVLMTGCVLLEKPNAGLVLTSTNAFFVIVKPIINEVPEESAMKWDVYL	MAVVASAPGKVLMTGCVLLEKPNAGLVLTSTNAFFVIVKPIINEVPEESAMKWDVYL	TSPLSRESMYKSLNHLTLOSASDSRNPFEVHAIOVIAAAHLATPEKDESLKLL	TSPLSRESMYKSLNHLTLOSASDSRNPFEVHAIOVIAAAHLATPEKDESLKLL
60	60	120	120

```

OY 121 OGDLITILGNDYFYRYNOIESAGLPLTPESLGTLPAPASTEFNAESNGANSKREPVAKT 180
DB 121 OGDLITILGNDYFYRYNOIESAGLPLTPESLGTLPAPASTEFNAESNGANSKREPVAKT 180
OY 181 GIGSSAAMTTAVAAALHLHGLVNDLSDPCKEKGFGSDLDVHIMIAOTSHCLAOQKVGSG 240
DB 181 GIGSSAAMTTAVAAALHLHGLVNDLSDPCKEKGFGSDLDVHIMIAOTSHCLAOQKVGSG 240
OY 241 PDVSCAVYGSQRYVRESPEVLSFAQVAVTGLPLNEVIGTILGKMDKNTREESLPLMLN 300
DB 241 PDVSCAVYGSQRYVRESPEVLSFAQVAVTGLPLNEVIGTILGKMDKNTREESLPLMLN 300
OY 301 FLGEPGSGSSSTPSPWGAVKKQMSDPEKARENMOMLSANLELETKLNDLSKLAKDHMD 360
DB 301 FLGEPGSGSSSTPSPWGAVKKQMSDPEKARENMOMLSANLELETKLNDLSKLAKDHMD 360
OY 361 VYLRVYKSCSVLTSEKVMVHATEPINEATIKELLEAREAMLRIRILMONGEAASVP1EP 420
DB 361 VYLRVYKSCSVLTSEKVMVHATEPINEATIKELLEAREAMLRIRILMONGEAASVP1EP 420
OY 421 ESOTOLDSTMSAEGVLLAGVPAGGFDAIFAITLGDSCGTLQAMSSHNVLALLVREDP 480
DB 421 ESOTOLDSTMSAEGVLLAGVPAGGFDAIFAITLGDSCGTLQAMSSHNVLALLVREDP 480
OY 481 HGVCLESQDPRTTCITSGVSSIHLE 505
DB 481 HGVCLESQDPRTTCITSGVSSIHLE 505

```

RESULT 2

```

PCT-US02-24048A-12
; Sequence 12, Application PC/TUS0224048A
; GENERAL INFORMATION:
; APPLICANT: E.I. du Pont de Nemours & Company
; TITLE OF INVENTION: Genes Involved in the Biosynthesis of Isopentenyl Diphosphate in
; FILE REFERENCE: CL1792 PCT
; CURRENT APPLICATION NUMBER: PCT/US02/24048A
; PRIOR FILING DATE: 2002-09-10
; PRIOR APPLICATION NUMBER: 60/307,637
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 12
; LENGTH: 503
; TYPE: PRT
; ORGANISM: Hevea brasiliensis
; PCT-US02-24048A-12

```

```

Query Match 74.8%; Score 1930.5; DB 1; Length 503;
Best Local Similarity 74.1%; Pred. No. 5.9e-182;
Matches 374; Conservative 61; Mismatches 67; Indels 3; Gaps 2;

```

```

OY 1 MAVVASAPGKVLMTGGYLVLEKPNAGLVLTSTNARFYAIKPTINEEYKPPESAMAKMTDVKL 60
DB 1 MAVVASAPGKVLMTGGYLVLEKPNAGLVLTSTNARFYAIKPTINEEYKPPESAMAKMTDVKL 60
OY 61 TSPQLSRESMYKLSLNHLTLQSVASDSRNPFEYEAIOYAIAAHATLATEKDKES-LHKLL 119
DB 61 TSPQLSRESMYKLSLNHLTLQSVASDSRNPFEYEAIOYAIAAHATLATEKDKES-LHKLL 119
OY 120 LOGDLITILGNDYFYRYNOIESAGLPLTPESLGTLPAPASTEFNAESNGANSKREPVAK 179
DB 120 LOGDLITILGNDYFYRYNOIESAGLPLTPESLGTLPAPASTEFNAESNGANSKREPVAK 179
OY 181 TGLSSAAMTTAVAAALHLHGLVNDLSDPCKEKGFGSDLDVHIMIAOTSHCLAOQKVGSG 238
DB 181 TGLSSAAMTTAVAAALHLHGLVNDLSDPCKEKGFGSDLDVHIMIAOTSHCLAOQKVGSG 238
OY 240 GPDVSAVYGSQRYVRESPEVLSFAQVAVTGLPLNEVIGTILGKMDKNTREESLPLMLN 299
DB 240 GPDVSAVYGSQRYVRESPEVLSFAQVAVTGLPLNEVIGTILGKMDKNTREESLPLMLN 299
OY 481 HGVCLESQDPRTTCITSGVSSIHLE 505
DB 481 HGVCLESQDPRTTCITSGVSSIHLE 505

```

```

OY 300 LFLGEPGSGSSSTPSPWGAVKKQMSDPEKARENMOMLSANLELETKLNDLSKLAKDHMD 359
DB 299 LFLGEPGSGSSSTPSPWGAVKKQMSDPEKARENMOMLSANLELETKLNDLSKLAKDHMD 358
OY 360 DYLVRVYKSCSVLTSEKVMVHATEPINEATIKELLEAREAMLRIRILMONGEAASVP1EP 419
DB 359 DYLVRVYKSCSVLTSEKVMVHATEPINEATIKELLEAREAMLRIRILMONGEAASVP1EP 418
OY 420 PESOTOLDSTMSAEGVLLAGVPAGGFDAIFAITLGDSCGTLQAMSSHNVLALLVREDP 479
DB 419 PESOTOLDSTMSAEGVLLAGVPAGGFDAIFAITLGDSCGTLQAMSSHNVLALLVREDP 478
OY 480 HGVCLESQDPRTTCITSGVSSIHLE 504
DB 479 PNGVLLSDDPRTTCITSGVSSIHLE 503

```

RESULT 3

```

US-10-036-959B-12
; Sequence 12, Application US/10036959B
; GENERAL INFORMATION:
; APPLICANT: E.I. du Pont de Nemours & Company
; TITLE OF INVENTION: Genes Involved in the Biosynthesis of Isopentenyl Diphosphate
; FILE REFERENCE: CL-1792
; CURRENT APPLICATION NUMBER: US-10/036,959B
; PRIOR FILING DATE: 2002-05-10
; PRIOR APPLICATION NUMBER: 60/307,637
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 12
; LENGTH: 503
; TYPE: PRT
; ORGANISM: Hevea brasiliensis
; US-10-036-959B-12

```

```

Query Match 74.8%; Score 1930.5; DB 24; Length 503;
Best Local Similarity 74.1%; Pred. No. 5.9e-182;
Matches 374; Conservative 61; Mismatches 67; Indels 3; Gaps 2;

```

```

OY 1 MAVVASAPGKVLMTGGYLVLEKPNAGLVLTSTNARFYAIKPTINEEYKPPESAMAKMTDVKL 60
DB 1 MAVVASAPGKVLMTGGYLVLEKPNAGLVLTSTNARFYAIKPTINEEYKPPESAMAKMTDVKL 60
OY 61 TSPQLSRESMYKLSLNHLTLQSVASDSRNPFEYEAIOYAIAAHATLATEKDKES-LHKLL 119
DB 61 TSPQLSRESMYKLSLNHLTLQSVASDSRNPFEYEAIOYAIAAHATLATEKDKES-LHKLL 119
OY 120 LOGDLITILGNDYFYRYNOIESAGLPLTPESLGTLPAPASTEFNAESNGANSKREPVAK 179
DB 120 LOGDLITILGNDYFYRYNOIESAGLPLTPESLGTLPAPASTEFNAESNGANSKREPVAK 179
OY 181 TGLSSAAMTTAVAAALHLHGLVNDLSDPCKEKGFGSDLDVHIMIAOTSHCLAOQKVGSG 238
DB 181 TGLSSAAMTTAVAAALHLHGLVNDLSDPCKEKGFGSDLDVHIMIAOTSHCLAOQKVGSG 238
OY 240 GPDVSAVYGSQRYVRESPEVLSFAQVAVTGLPLNEVIGTILGKMDKNTREESLPLMLN 299
DB 240 GPDVSAVYGSQRYVRESPEVLSFAQVAVTGLPLNEVIGTILGKMDKNTREESLPLMLN 299
OY 300 LFLGEPGSGSSSTPSPWGAVKKQMSDPEKARENMOMLSANLELETKLNDLSKLAKDHMD 359
DB 299 LFLGEPGSGSSSTPSPWGAVKKQMSDPEKARENMOMLSANLELETKLNDLSKLAKDHMD 358
OY 420 PESOTOLDSTMSAEGVLLAGVPAGGFDAIFAITLGDSCGTLQAMSSHNVLALLVREDP 479
DB 419 PESOTOLDSTMSAEGVLLAGVPAGGFDAIFAITLGDSCGTLQAMSSHNVLALLVREDP 478

```

Db 419 PESQRLDITMNDGVLLAGVPGAGFDVAFVATLGDSTGNVAKWSSLNALLVRED 478
 Oy 480 PHGVLESQDPRTCITSGVSSIH. 504
 Db 479 PNGVLESQDPRTCITTAFAVAHI 503

RESULT 4 US-10-219-999-40847

; Sequence 40847: Application US/10219999
 ; GENERAL INFORMATION:
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Edgerton, Michael D
 ; APPLICANT: Hinkle, Gregory J.
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Liu, Jindong
 ; APPLICANT: Stein, Joshua
 ; TITLE OF INVENTION: CDNA SEQUENCES AND USES FOR PLANT IMPROVEMENT
 ; FILE REFERENCE: 38-10(52726)C
 ; CURRENT APPLICATION NUMBER: US/10/219,999
 ; CURRENT FILING DATE: 2002-08-15
 ; PRIOR APPLICATION NUMBER: US 60/324,109
 ; PRIOR FILING DATE: 2001-09-21
 ; PRIOR APPLICATION NUMBER: US 60/312,544
 ; PRIOR FILING DATE: 2001-08-15
 ; NUMBER OF SEQ ID NOS: 63520
 ; SEQ ID NO 40847
 ; LENGTH: 511
 ; TYPE: PR1
 ; ORGANISM: Zea mays
 US-10-219-999-40847

Query Match 61.2%; Score 1579; DB 26; Length 511;
 Best local Similarity 62.6%; Pred. No. 5.7e-147;
 Matches 321; Conservative 69; Mismatches 111; Indels 12; Gaps 5;

Oy 1 MAVVASAPGKVLMTGTYLEKPNAGLVLTSTNARFYAIRPINEVPEKSNAMKMTDVKL 60
 Db 1 MEVVASAPGKVLMTGTYLEKPNAGLVLTSTNARFYAIRPINEVPEKSNAMKMTDVKL 60
 Oy 61 TSPQLSEMYKLSLNHLTQSVASDSRNPFEVHAIOYIAAAHLATEKDKES---LHK 117
 Db 61 TSPQLSEMYKLSLNHLTQSVASDSRNPFEVHAIOYIAAAHLATEKDKES---LHK 117
 Oy 118 LLOGLDITLIGSNDFTSYRNOIESAGLPLTPRESLGTLPAPASTTFNNAESNGA----NS 173
 Db 120 LLOGLDITLIGSNDFTSYRNOIESAGLPLTPRESLGTLPAPASTTFNNAESNGA----NS 173
 Oy 174 KPEVAKTGLGSSAAMTAVVAALLHYLVGVDLSDPCKEGRGCS--DLDVDIHMIAQTSHC 231
 Db 180 KPEVAKTGLGSSAAMTAVVAALLHYLVGVDLSDPCKEGRGCS--DLDVDIHMIAQTSHC 231
 Oy 232 LAQKVGSGDFVSCAVYGSQRYVRFSPFVLSFAQVAVTGLPLNEVIGTILKGMKDNKTE 291
 Db 240 LAQKVGSGDFVSCAVYGSQRYVRFSPFVLSFAQVAVTGLPLNEVIGTILKGMKDNKTE 291
 Oy 292 FSLPPLNMLFLGEBGSGSTSPSWGAVKKQMSDPEKARQMONLS DANLEETKLNLD. 351
 Db 298 FSLPPLNMLFLGEBGSGSTSPSWGAVKKQMSDPEKARQMONLS DANLEETKLNLD. 351
 Oy 352 SKLAKDHMDVLYRIKSCSVLTSEKWLHATEPINEAIIKELLEAREAMLRIRILRMONG 411
 Db 358 KGLSENHREAVESVYLSCHLTYEKTWEVATNOHOELIKSLAARDACLEIRLHMREMG 417
 Oy 412 EAASVPIEPESQOTLDLSTMSAEGVLLAGVPGAGFDVAFVATLGDSTGKLTQAMSSHN 471
 Db 418 IAAGVPIEPESQOTLDLSTMSAEGVLLAGVPGAGFDVAFVATLGDSTGKLTQAMSSHN 471
 Oy 472 LALLVREDPHVCLESQDPRTCITSGVSSIH. 504
 Db 478 LPLLVREDPHVCLESQDPRTCITSGVSSIH. 510

RESULT 5 US-60-324-109-16899

; Sequence 16899: Application US/603241095
 ; GENERAL INFORMATION:
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Edgerton, Michael D
 ; APPLICANT: Hinkle, Gregory J.
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Liu, Jindong
 ; APPLICANT: Stein, Joshua
 ; TITLE OF INVENTION: CDNA SEQUENCES AND USES FOR PLANT IMPROVEMENT
 ; FILE REFERENCE: 38-10(52726)B
 ; CURRENT APPLICATION NUMBER: US/60/324,109
 ; CURRENT FILING DATE: 2001-09-21
 ; NUMBER OF SEQ ID NOS: 33196
 ; SEQ ID NO 16899
 ; LENGTH: 511
 ; TYPE: PR1
 ; ORGANISM: Zea mays
 ; FEATURE:
 US-60-324-109-16899

Query Match 61.2%; Score 1579; DB 27; Length 511;
 Best local Similarity 62.6%; Pred. No. 5.7e-147;
 Matches 321; Conservative 69; Mismatches 111; Indels 12; Gaps 5;

Oy 1 MAVVASAPGKVLMTGTYLEKPNAGLVLTSTNARFYAIRPINEVPEKSNAMKMTDVKL 60
 Db 1 MEVVASAPGKVLMTGTYLEKPNAGLVLTSTNARFYAIRPINEVPEKSNAMKMTDVKL 60
 Oy 61 TSPQLSEMYKLSLNHLTQSVASDSRNPFEVHAIOYIAAAHLATEKDKES---LHK 117
 Db 61 TSPQLSEMYKLSLNHLTQSVASDSRNPFEVHAIOYIAAAHLATEKDKES---LHK 117
 Oy 118 LLOGLDITLIGSNDFTSYRNOIESAGLPLTPRESLGTLPAPASTTFNNAESNGA----NS 173
 Db 120 LLOGLDITLIGSNDFTSYRNOIESAGLPLTPRESLGTLPAPASTTFNNAESNGA----NS 173
 Oy 174 KPEVAKTGLGSSAAMTAVVAALLHYLVGVDLSDPCKEGRGCS--DLDVDIHMIAQTSHC 231
 Db 180 KPEVAKTGLGSSAAMTAVVAALLHYLVGVDLSDPCKEGRGCS--DLDVDIHMIAQTSHC 231
 Oy 232 LAQKVGSGDFVSCAVYGSQRYVRFSPFVLSFAQVAVTGLPLNEVIGTILKGMKDNKTE 291
 Db 240 LAQKVGSGDFVSCAVYGSQRYVRFSPFVLSFAQVAVTGLPLNEVIGTILKGMKDNKTE 291
 Oy 292 FSLPPLNMLFLGEBGSGSTSPSWGAVKKQMSDPEKARQMONLS DANLEETKLNLD. 351
 Db 298 FSLPPLNMLFLGEBGSGSTSPSWGAVKKQMSDPEKARQMONLS DANLEETKLNLD. 351
 Oy 352 SKLAKDHMDVLYRIKSCSVLTSEKWLHATEPINEAIIKELLEAREAMLRIRILRMONG 411
 Db 358 KGLSENHREAVESVYLSCHLTYEKTWEVATNOHOELIKSLAARDACLEIRLHMREMG 417
 Oy 412 EAASVPIEPESQOTLDLSTMSAEGVLLAGVPGAGFDVAFVATLGDSTGKLTQAMSSHN 471
 Db 418 IAAGVPIEPESQOTLDLSTMSAEGVLLAGVPGAGFDVAFVATLGDSTGKLTQAMSSHN 471
 Oy 472 LALLVREDPHVCLESQDPRTCITSGVSSIH. 504
 Db 478 LPLLVREDPHVCLESQDPRTCITSGVSSIH. 510

RESULT 6

US-09-708-427-83408
 ; Sequence 83408: Application US/09708427
 ; GENERAL INFORMATION:
 ; APPLICANT: N. ALEXANDROV et al.
 ; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID
 ; FILE REFERENCE: 2750-1243P
 ; CURRENT APPLICATION NUMBER: US/09/708,427
 ; CURRENT FILING DATE: 2000-11-09

```

; NUMBER OF SEQ ID NOS: 85364
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 83408
; LENGTH: 499
; TYPE: PRF
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 1..499
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc.feature
; LOCATION: 1..499
; OTHER INFORMATION: Ceres Seq. ID 1962605
US-09-708-427-83408

```

```

Query Match          59.3%; Score 1531; DB 21; Length 499;
Best Local Similarity 61.0%; Pred. No. 3.2e-142;
Matches 313; Conservative 68; Mismatches 108; Indels 24; Gaps 6;

```

```

QY 1 MAVVASAPGKVLMTGGLVLEKPNAGLVSTNARFVAIVPINEVPEESMAKMTDVKL 60
  1 MAVVASAPGKVLMTGGLVLEKPNAGLVSTNARFVAIVPINEVPEESMAKMTDVKV 60
Db 1 MEVASAPGKVLMTGGLVLEKPNAGLVSTNARFVAIVPINEVPEESMAKMTDVKV 60
QY 61 TSPQLSESMYKLSLNLTLQSVASDSRNPVEHAIQVIAAAHLATEKDKES---LHK 117
  61 TSPQLSESMYKLSLNLTLQSVASDSRNPVEHAIQVIAAAHLATEKDKES---LHK 117
Db 61 TSPQLSESMYKLSLNLTLQSVASDSRNPVEHAIQVIAAAHLATEKDKES---LHK 117
QY 118 LLLQGLDITLLGNDPFSYRNOIESAGLPLTPESLGLAPFASITFNAESNGA---NS 173
  118 LLLQGLDITLLGNDPFSYRNOIESAGLPLTPESLGLAPFASITFNAESNGA---NS 173
Db 120 LLLQGLDITLLGNDPFSYRNOIESAGLPLTPESLGLAPFASITFNAESNGA---NS 173
QY 174 KPEVAKTGLSSAAMTAVVAALLHYLVGYVLDSDPCKEGFGCS--DLVDYIHMIAQTSHC 231
  174 KPEVAKTGLSSAAMTAVVAALLHYLVGYVLDSDPCKEGFGCS--DLVDYIHMIAQTSHC 231
Db 180 KPEVAKTGLSSAAMTAVVAALLHYLVGYVLDSDPCKEGFGCS--DLVDYIHMIAQTSHC 231
QY 232 LAGKVGSGFDVSAAYGSGRYRFPSEVLSPFQAVATGLPNEVIGTILKGKMDNKRT 291
  232 LAGKVGSGFDVSAAYGSGRYRFPSEVLSPFQAVATGLPNEVIGTILKGKMDNKRT 291
Db 240 LAGKVGSGFDVSAAYGSGRYRFPSEVLSPFQAVATGLPNEVIGTILKGKMDNKRT 291
QY 292 FSLPPLMTLLGEGTGSSTPSWVGAVKRWKMSDPKAKENQNLSDANILETKLNDL 351
  292 FSLPPLMTLLGEGTGSSTPSWVGAVKRWKMSDPKAKENQNLSDANILETKLNDL 351
Db 298 FSLPPLMTLLGEGTGSSTPSWVGAVKRWKMSDPKAKENQNLSDANILETKLNDL 351
QY 352 SKLAKHDVYLVKRYKCSVLTSKRVLAHTEPINEAIIELLEPARAMLRIRILMONG 411
  352 SKLAKHDVYLVKRYKCSVLTSKRVLAHTEPINEAIIELLEPARAMLRIRILMONG 411
Db 358 KGLSEHREAY-----ESWTEVATNQHOLEIISLLAARACLEIRLHREMKG 405
QY 412 EAASVPIEPSOTQLDSTWASGVLGAGVPGAGGPAIFAITLGDGCTKLTOAMSSHN 471
  412 EAASVPIEPSOTQLDSTWASGVLGAGVPGAGGPAIFAITLGDGCTKLTOAMSSHN 471
Db 406 IAGVPIEPSOTQLDSTWASGVLGAGVPGAGGPAIFAITLGDGCTKLTOAMSSHN 471
QY 472 LALLVREDPHGVCLSGDPRPTTCITSGVSSIHL 504
  472 LALLVREDPHGVCLSGDPRPTTCITSGVSSIHL 504
Db 466 LPLLVREDCGVSLDADPRTREVSAAVSSIQI 498

```

```

RESULT 7
US-09-708-427-83407
; Sequence 83407, Application US/09708427
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1243P
; CURRENT APPLICATION NUMBER: US/09/708,427
; CURRENT FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 85364
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 83407
; LENGTH: 535
; TYPE: PRF
; ORGANISM: Zea mays subsp. mays
; FEATURE:

```

```

; NAME/KEY: misc.feature
; LOCATION: 1..535
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc.feature
; LOCATION: 1..535
; OTHER INFORMATION: Ceres Seq. ID 1962604
US-09-708-427-83407

```

```

Query Match          59.3%; Score 1531; DB 21; Length 535;
Best Local Similarity 61.0%; Pred. No. 3.7e-142;
Matches 313; Conservative 68; Mismatches 108; Indels 24; Gaps 6;

```

```

QY 1 MAVVASAPGKVLMTGGLVLEKPNAGLVSTNARFVAIVPINEVPEESMAKMTDVKL 60
  1 MAVVASAPGKVLMTGGLVLEKPNAGLVSTNARFVAIVPINEVPEESMAKMTDVKV 60
Db 37 MEVASAPGKVLMTGGLVLEKPNAGLVSTNARFVAIVPINEVPEESMAKMTDVKV 96
QY 61 TSPQLSESMYKLSLNLTLQSVASDSRNPVEHAIQVIAAAHLATEKDKES---LHK 117
  61 TSPQLSESMYKLSLNLTLQSVASDSRNPVEHAIQVIAAAHLATEKDKES---LHK 117
Db 97 TSPQLSESMYKLSLNLTLQSVASDSRNPVEHAIQVIAAAHLATEKDKES---LHK 155
QY 118 LLLQGLDITLLGNDPFSYRNOIESAGLPLTPESLGLAPFASITFNAESNGA---NS 173
  118 LLLQGLDITLLGNDPFSYRNOIESAGLPLTPESLGLAPFASITFNAESNGA---NS 173
Db 156 LLLQGLDITLLGNDPFSYRNOIESAGLPLTPESLGLAPFASITFNAESNGA---NS 215
QY 174 KPEVAKTGLSSAAMTAVVAALLHYLVGYVLDSDPCKEGFGCS--DLVDYIHMIAQTSHC 231
  174 KPEVAKTGLSSAAMTAVVAALLHYLVGYVLDSDPCKEGFGCS--DLVDYIHMIAQTSHC 231
Db 216 KPEVAKTGLSSAAMTAVVAALLHYLVGYVLDSDPCKEGFGCS--DLVDYIHMIAQTSHC 275
QY 232 LAGKVGSGFDVSAAYGSGRYRFPSEVLSPFQAVATGLPNEVIGTILKGKMDNKRT 291
  232 LAGKVGSGFDVSAAYGSGRYRFPSEVLSPFQAVATGLPNEVIGTILKGKMDNKRT 291
Db 276 LAGKVGSGFDVSAAYGSGRYRFPSEVLSPFQAVATGLPNEVIGTILKGKMDNKRT 333
QY 292 FSLPPLMTLLGEGTGSSTPSWVGAVKRWKMSDPKAKENQNLSDANILETKLNDL 351
  292 FSLPPLMTLLGEGTGSSTPSWVGAVKRWKMSDPKAKENQNLSDANILETKLNDL 351
Db 334 FSLPPLMTLLGEGTGSSTPSWVGAVKRWKMSDPKAKENQNLSDANILETKLNDL 393
QY 352 SKLAKHDVYLVKRYKCSVLTSKRVLAHTEPINEAIIELLEPARAMLRIRILMONG 411
  352 SKLAKHDVYLVKRYKCSVLTSKRVLAHTEPINEAIIELLEPARAMLRIRILMONG 411
Db 394 KGLSEHREAY-----ESWTEVATNQHOLEIISLLAARACLEIRLHREMKG 441
QY 412 EAASVPIEPSOTQLDSTWASGVLGAGVPGAGGPAIFAITLGDGCTKLTOAMSSHN 471
  412 EAASVPIEPSOTQLDSTWASGVLGAGVPGAGGPAIFAITLGDGCTKLTOAMSSHN 471
Db 442 IAGVPIEPSOTQLDSTWASGVLGAGVPGAGGPAIFAITLGDGCTKLTOAMSSHN 501
QY 472 LALLVREDPHGVCLSGDPRPTTCITSGVSSIHL 504
  472 LALLVREDPHGVCLSGDPRPTTCITSGVSSIHL 504
Db 502 LPLLVREDCGVSLDADPRTREVSAAVSSIQI 534

```

```

RESULT 8
US-09-708-427-66289
; Sequence 66289, Application US/09708427
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID
; FILE REFERENCE: 2750-1243P
; CURRENT APPLICATION NUMBER: US/09/708,427
; CURRENT FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 85364
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 66289
; LENGTH: 535
; TYPE: PRF
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 1..535
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc.feature
; LOCATION: 1..535
; OTHER INFORMATION: Ceres Seq. ID 1930323
US-09-708-427-66289

```

Query Match	52.0%;	Score 1341;	DB 21;	Length 535;
Best Local Similarity	59.6%;	Pred. No. 2.9e-123;		
Matches 277; Conservative	61;	Mismatches 103;	Indels 24;	Gaps 6

[illegible]

```

RESULT 9
US-09-708-427-66290
; Sequence 66290, Application US/09708427
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE OF INVENTION: THERREBY
; FILE REFERENCE: 2750-1243P
; CURRENT APPLICATION NUMBER: US/09/708,427
; CURRENT FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 85364
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 66290
; LENGTH: 326
; TYPE: PRT
; ORGANISM: Zea mays subsp. mays
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..326
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: 1..326
; OTHER INFORMATION: Ceres Seq. ID 1930324
; US-09-708-427-66290

```

```

Query Match      37.7%  Score 973  DB 21  Length 326
Best Local Similarity 58.8%  Pred. No. 5,1e-87
Matches 198; Conservative 45; Mismatches 78; Indels 16; Gaps 3.

QY .170 GANSKPEYAKTGLGSSAMTTAVVAALILHYLGVDLSPDCKEKGKFGCS--DEDVHIMIAO 227
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db  3  GKCKPEYAKTGLGSSAMTTGSVAALLHYLGVDLSPDCKEKGKFGCSGNTTRREDIVHISMD 62
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 228 TSHCLAQKGVSGFVSCAVYVGSQKRYRSPVLSFAOVAYATGDLPLNEYIGTILKGMKN 287
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

Db      63  SAHCAQCKISSGDFVSAANYGGRVRYRFPSEILLISQATGGTFLP--DVVSDIVTQKMDH 120
OY      288  KRTESELPPLMNLFLGEPGSGSSPTSMGAVKMKQMSDPEKARENNQNSDANILETK 347
       : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      121  ENKFOSTPLMTLLLGEGCTGSSPTSMGVSKKMKQSDPECKCDTMSKLAFANSLENO 180
OY      348  LNDISKLAKHDWVLYLRVYKSCSVLTSEKWLHATPEINFAITIKELLEAREAMLRITLM 407
       : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      181  LRLLKGLSENIREFX-----ESMTFVATNQHOLEITIKSLAAMDACLIEIRLHM 228
OY      408  RQMGEPASVPIEPESQTLDDSTNSAEGCVLLAGVPGAGGFDALFAITLIGDSGTRKLTQWMS 467
       : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      229  REMGIAGVPIEPESQTRILDATNMKGCVLLAGVPGAGGFDVAFSVYTLGDASNAVANAMS 288
OY      468  SHNVATILVREDPHGVCLSESDPPTCTTICTSVSYIHL 504
       : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      289  SAGVLPFLVREDCKRVSLEDDATREVSAAVSSIQI 325

```

```

RESULT 10
US-09-708-427-83409
: Sequence 83409, Application US/09708427
:
: GENERAL INFORMATION:
: APPLICANT: N. ALEXANDROV et al.
: TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID
: TITLE OF INVENTION: THEREBY
: FILE REFERENCE: 2750-1243P
: CURRENT APPLICATION NUMBER: US/09/708,427
: CURRENT FILING DATE: 2000-11-09
: NUMBER OF SEQ ID NOS: 85364
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 83409
: LENGTH: 326
: TYPE: prt
: ORGANISM: Zea mays subsp. mays
:
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: 1..326
: OTHER INFORMATION: Xaa is any amino acid
: NAME/KEY: misc_feature
: LOCATION: 1..326
: OTHER INFORMATION: Ceres Seq. ID 1962606
US-09-708-427-83409

```

[illegible]

Query Match	34.9%	Score 902:	DB 21:	Length 306;
Best Local Similarity	57.4%;	Pred No. 5.2e-80;		
Matches 183;	Conservative 45;	Mismatches 75;	Indels 16;	Gaps 3

RESULT 12
US-10-219-999-55015

```

: SEQ ID NO 55015
: LENGTH: 286
: TYPE: PRT
: ORGANISM: Zea mays
US-10-219-999-55015

Query Match
Best Local Similarity 33.6%; Score 868.5; DB 26; Length 286;
Matches 170; Conservative 45; Mismatches 71; Indels 1; Gaps 1.

Oy 218 DLDVHTHIAQTSHCLAQGVSGFPEVSCAVYGSQRYVRFPSPVLSFAQVAAYVGLPLNEVI 277
Db 1 ELDIYHAIQAOSHCLAQGVSGFPEVSAAYVGSQRYVRFPSPVLSAQA-AIGTVLPDVV 59
Oy 278 GTILKGMKDNKRTFESPLPLMLPLFGPEPGSGSSPPSMWGAAYKKMOMSPPEKARENMONTL 337
Db 60 SDVLTQRNDHEKQPSLPLMLTLGEGPTGSGSSPPSMWGSYKRWLKSPEKSRDTWSKL 119
Oy 338 SDANLEETKLANDLSKLAKDHDVYLRYIKSCSVLTSEKWLHATEPINEALIKELLEAR 397
Db 120 AIASTLTENQRIKLGLSENHHEAVESWYRSCSRLLTGYKMAEVAANOHELLIRSLAAR 179
Oy 398 EAMLRIRIMOMGEAASVPIPEESOTQLDSTMSAEGVLLAGVPGAGGFDALFAITLGD 457
Db 180 NACLEFRLHMRPMGLAAGVPIPEDSQTRLIDATMNEGVLLAGVPGAGGFDALFVSVLGD 239
Oy 458 SGTKLTOAMSSHNVLALLVREDPHGVCLSESGDPRITTCITSGVSIHL 504
Db 240 ASNAVAHAMSSAGVLPPLVREDRCGVSLSDADDPRTREVSAAVMSIQI 286

RESULT 13
US-10-219-999-42346
: Sequence 42346, Application US/10219999
: GENERAL INFORMATION:
: APPLICANT: Cao, Yongwei
: APPLICANT: Edgerton, Michael D
: APPLICANT: Hinkle, Gregory J.
: APPLICANT: Kovalic, David K.
: APPLICANT: Liu, Jingdong
: APPLICANT: Stein, Joshua
: TITLE OF INVENTION: cDNA SEQUENCES AND USES FOR PLANT IMPROVEMENT
: FILE REFERENCE: 38-10(52726)C
: CURRENT APPLICATION NUMBER: US/10/219, 999
: CURRENT FILING DATE: 2002-08-15
: PRIOR APPLICATION NUMBER: US 60/324,109
: PRIOR FILING DATE: 2001-09-21
: PRIOR APPLICATION NUMBER: US 60/312,544
: PRIOR FILING DATE: 2001-08-15
: NUMBER OF SEQ ID NOS: 63520
: SEQ ID NO 42346
: LENGTH: 281
: TYPE: PRT
: ORGANISM: Zea mays
US-10-219-999-42346

Query Match
Best Local Similarity 32.5%; Score 839.5; DB 26; Length 281;
Matches 166; Conservative 43; Mismatches 70; Indels 1; Gaps 1.

Oy 225 IAOTSHCLAQGVSGFDVSCAVYGSQRYVRFPSEYLSFAQVAAYVGLPLNEVIYGTILKGR 284
Db 2 IAOSHCHCLAQGVSGFDVSAAYVGSQRYVRFPSPVLSAQA-AIGTVLPDVVSDVLTQR 60
Oy 285 WDNKRTFESPLPLMLPLFGPEPGSGSSPPSMWGAAYKKMOMSPPEKARENMONTLSDANLE 344
Db 61 WDHENKQPSLPLMLTLGEGPTGSGSSPPSMWGSYKRWLKSPEKSRDTWSKLALANSTL 120
Oy 345 EFKLNDLSKLAKDHDVYLRYIKSCSVLTSEKWLHATEPINEALIKELLEAREAMLRIR 404
Db 121 ENQALFRLHMRPMGLAAGVPIPEDSQTRLIDATMNEGVLLAGVPGAGGFDALFVSVLGD 180
Oy 405 IIMROMGEAASVPIPEESOTQLDSTMSAEGVLLAGVPGAGGFDALFAITLGDSGTKLQ 464

```


Db 181 LHMREMGIAAGVIEPDSQTRLLDATMNMGGVLLAGVPGAGFDVAFSVLLGDASNAVAH 240
 QY 465 AMSSHNVALLVREDPHGVCLSGDPRCTCITSGVSSIHL 504
 Db 241 AMSSVGVLLPLVREDCRGVSLDADPRTRVSAVWSIQI 280

RESULT 14
 US-60-312-544-6223

; Sequence 6223, Application US/60312544
 ; GENERAL INFORMATION:
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Edgerton, Michael D
 ; APPLICANT: Hinkle, Gregory J.
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Liu, Jingdong
 ; APPLICANT: Stein, Joshua
 ; TITLE OF INVENTION: CDNA SEQUENCES AND USES FOR PLANT IMPROVEMENT
 ; FILE REFERENCE: 38-10(52726)A
 ; CURRENT APPLICATION NUMBER: US/60/312,544
 ; CURRENT FILING DATE: 2001-08-15
 ; NUMBER OF SEQ ID NOS: 10730
 ; SEQ ID NO 6223
 ; LENGTH: 281
 ; TYPE: PRT
 ; ORGANISM: Zea mays
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: 700571779_FLI
 US-60-312-544-6223

Query Match 32.5%; Score 839.5; DB 27; Length 281;
 Best Local Similarity 59.3%; Pred. No. 7, 5e-74;
 Matches 166; Conservative 43; Mismatches 70; Indels 1; Gaps 1;

QY 225 IAGTSHCLAQKVGSGFDVSCAVYGSQRYRSPDEVLSFAQVAVTGILPNEVIGTILKGR 284
 Db 2 IAGSAHCLAQKIGSGFDVSAVYGSQRYRSPDEVLSAQ-AIGGTVPDVPVSDVLTQR 60
 QY 285 WDKKREFSLPPLMNLFLGPGSGSSPSMGAVKWKQMSDEPKARENNONSDALEL 344
 Db 61 WDHENKQFSLPPLMTLLGLGPGSSPSMGVSKRWMLKSDPEKSRDTWSKLAIANSTL 120
 QY 345 ETKLNDLSKLAKDMVDYLLVIRKSVLTSEKWLHATEPINEALIKELAREAMLRIR 404
 Db 121 ENGLRLIKGLSENHHEAYEMVSCSRITGKMAEVAATNOHELIIRSLAARDACLEIR 180
 QY 405 ILMRONGEAASVPIDEPDSQTLLDSTMSAGVLLAGVPGAGFDVAFVAILGDSGTILQ 464
 Db 181 LHMREMGIAAGVIEPDSQTRLLDATMNMGGVLLAGVPGAGFDVAFSVLLGDASNAVAH 240
 QY 465 AMSSHNVALLVREDPHGVCLSGDPRCTCITSGVSSIHL 504
 Db 241 AMSSVGVLLPLVREDCRGVSLDADPRTRVSAVWSIQI 280

RESULT 15

US-09-708-427-65887
 ; Sequence 65887, Application US/09708427
 ; GENERAL INFORMATION:
 ; APPLICANT: N. ALEXANDROV et al.
 ; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
 ; FILE REFERENCE: 2750-1243P
 ; CURRENT APPLICATION NUMBER: US/09/708,427
 ; CURRENT FILING DATE: 2000-11-09
 ; NUMBER OF SEQ ID NOS: 85364
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 65887
 ; LENGTH: 359
 ; TYPE: PRT
 ; ORGANISM: Zea mays subsp. mays
 ; FEATURE:

; NAME/KEY: misc_feature
 ; LOCATION: 1..359
 ; OTHER INFORMATION: Xaa is any amino acid
 ; NAME/KEY: misc_feature
 ; LOCATION: 1..359
 ; OTHER INFORMATION: Ceres Seq. ID 1929648
 US-09-708-427-65887

Query Match 28.0%; Score 722.5; DB 21; Length 359;
 Best Local Similarity 66.2%; Pred. No. 5e-62;
 Matches 133; Conservative 26; Mismatches 41; Indels 11; Gaps 5;

QY 49 ESWAMKTDVKTSPQLSRSEMYKLSLNHLTQSVASDSRNPVEHAIOYAIANAHLAT 108
 Db 114 DSWTMAWTDVKTSPQLSRVATYKLSLNTKTLLQTSRESNPNVEQAIQFSVAAR-AT 172
 QY 109 EKDKES---LHKLLOGDITILGSDNPFYSRNOIESAGLPPTRESIGTIAPFASITFNA 165
 Db 173 IIDKERDQVVDKLLLOGINLITIGHNDFYSRKQIEARGLPPTREVLISLPFESSITFNS 232
 QY 166 AESNGA---NSKPEVAKTIGSSAAMTTAVVALLHYLGVDLSDPCDKGKFGCS--DL 219
 Db 233 EVANGTGTGCKPEVAKTIGSSAAMTTAVVALLHYLGAVNLSCFGQSSGDNASGREL 292
 QY 220 DVYHMAQTSCLAQKVGSGFDVSCAVYGSQRYRSPDEVLSFAQVAVTG 270
 Db 293 DIVHTINQSAHCLAQKIGSGFDVSAVYGSQRYRSPDEVLSAQ-AIGG 342

Search completed: April 26, 2003, 13:00:27
 Job time: 151 secs

THIS PAGE BLANK (user)

OY 123 LDITLGSNDFFSYRNOIE-----SAGLPTLPESLGLTAPFASITRNAASNGANSKP 175
DB 183 LDVVVFADNDFYSOREQVAMIFELTSLTPRISLPLPPTPL-----PRPIP 233
OY 176 EVAKTGLSSAMTAVVAALHLYGVVDLSDPCKEKGKCGSDLVYHIMIAOTSCLAOG 235
DB 234 ATNKTGLSSAALVSLVSGLSLHITHSS---PEGSIADDDKAVIHSVADLAHCQAOG 290
OY 236 KVGSGEDVSCAVYGSQRYRFSPEVLS-FAQYAVTGLPLNE-----VIGTILGKMDN 287
DB 291 KVGSGEDVSCAVYGSHTLRTFSPSILTPMLSLAPSRPRPGSGTSPLLDALHPSKMDN 350
OY 288 KTEFSLPLMLNLFCEPESGSSSTPSMWGAYKKQMSPEKAREKQWLSANLELETK 347
DB 351 KSIPRLPKHLRLADY-SCGTDPPSFVSVLKMNRNDRKADVWALDANRALGEV 409
OY 348 LNDLSKLAKDHDVYLRVSKSVLTSEKVLH-ATEPINEAIKELREARAMLRIIL 406
DB 410 LNDM-----LNLPTTSPSRTLXHLRLASXXSIRAL 443
OY 407 MRONGEASVPIEPSQQLDSTMSAEGVLLAGVPAGGDAIFAITL 445
DB 444 LREMSDLSGVPIEPKQRTILDACQGVKGVGAGYDALYLAI 492

RESULT 2
US-10-320-797-3218

; Sequence 3218, Application US/10320797
; GENERAL INFORMATION:
; APPLICANT: Eroskin, Alexey M.
; APPLICANT: Zamudio, Carlos
; TITLE OF INVENTION: IDENTIFICATION OF ESSENTIAL GENES OF CRYPTOCOCCUS NEOFORMANS AND
; TITLE OF INVENTION: METHODS OF USE
; FILE REFERENCE: 10182-021-999
; CURRENT APPLICATION NUMBER: US/10/320,797
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: 60/341,261
; PRIOR FILING DATE: 2001-12-17
; NUMBER OF SEQ ID NOS: 3361
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3218
; LENGTH: 501
; TYPE: PRT
; ORGANISM: Cryptococcus neoformans
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (435)..(435)
; OTHER INFORMATION: Xaa - any amino acid
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (437)..(438)
; OTHER INFORMATION: Xaa - any amino acid
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (429)..(429)
; OTHER INFORMATION: Xaa - any amino acid
US-10-320-797-3218

Query Match 23.7%; Score 612.5; DB 6; Length 501;
Best Local Similarity 31.8%; Pred. No. 5.4e-47;
Matches 168; Conservative 75; Mismatches 165; Indels 121; Gaps 16;

OY 4 VASAPGKVLMTGYLVLEKPNAGLVLTNARFYAIVK-----PINEYV 46
DB 8 VVSPGKVLVLAGYLVLDHYSGLVIGTSSRFYSCVSRATSSRATSPSIDADINNDI 67
OY 47 KESNAWKTDVKTSPOLSRBS---MKLSLNHLTLQSVASD-----87
DB 68 DTDITQNKAT-LSVRAGDFPSDASTWVCISKPSASASAGEDGDEGSDGLYLKL 126
OY 88 -----SRNPFV-----EHAIOVAIAAAHLATEKDKSILHLLD-----G 122

DB 127 EOTNEQAGKKNKFIITLCKVLEYV--YESILAHVGD--ETALDELKMKRIKGGGCGG 182
OY 123 LDITLGSNDFFSYRNOIE-----SAGLPTLPESLGLTAPFASITRNAASNGANSKP 175
DB 183 LDVVVFADNDFYSOREQVAMIFELTSLTPRISLPLPPTPL-----PRPIP 233
OY 176 EVAKTGLSSAMTAVVAALHLYGVVDLSDPCKEKGKCGSDLVYHIMIAOTSCLAOG 235
DB 234 ATNKTGLSSAALVSLVSGLSLHITHSS---PEGSIADDDKAVIHSVADLAHCQAOG 290
OY 236 KVGSGEDVSCAVYGSQRYRFSPEVLS-FAQYAVTGLPLNE-----VIGTILGKMDN 287
DB 291 KVGSGEDVSCAVYGSHTLRTFSPSILTPMLSLAPSRPRPGSGTSPLLDALHPSKMDN 350
OY 288 KTEFSLPLMLNLFCEPESGSSSTPSMWGAYKKQMSPEKAREKQWLSANLELETK 347
DB 351 KSIPRLPKHLRLADY-SCGTDPPSFVSVLKMNRNDRKADVWALDANRALGEV 409
OY 348 LNDLSKLAKDHDVYLRVSKSVLTSEKVLH-ATEPINEAIKELREARAMLRIIL 406
DB 410 LNDM-----LNLPTTSPSRTLXHLRLASXXSIRAL 443
OY 407 MRONGEASVPIEPSQQLDSTMSAEGVLLAGVPAGGDAIFAITL 445
DB 444 LREMSDLSGVPIEPKQRTILDACQGVKGVGAGYDALYLAI 492

RESULT 3
US-09-909-745A-22

; Sequence 22, Application US/09909745A
; GENERAL INFORMATION:
; APPLICANT: Falco, S. Carl
; APPLICANT: Pamodu, Onciayo O.
; TITLE OF INVENTION: Squalene Synthase Enzymes
; FILE REFERENCE: B01112 US CIP
; CURRENT APPLICATION NUMBER: US/09/909,745A
; CURRENT FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: 09/433,242
; PRIOR FILING DATE: 1999-11-04
; PRIOR APPLICATION NUMBER: 60/107,241
; PRIOR FILING DATE: 1998-11-05
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 22
; LENGTH: 179
; TYPE: PRT
; ORGANISM: Oryza sativa
US-09-909-745A-22

Query Match 22.3%; Score 575; DB 5; Length 179;
Best Local Similarity 62.0%; Pred. No. 2.5e-44;
Matches 111; Conservative 26; Mismatches 42; Indels 0; Gaps 0;

OY 308 GGSSTPSMWGAYKKQMSPEKAREKQWLSANLELETKLNDLSKLAKDHDVYLRVYK 367
DB 1 GGSSTPSMWGAYKKQMSPEKAREKQWLSANLELETKLNDLSKLAKDHDVYLRVYK 60
OY 368 SCVSILTSEKVLHATEPINEAIKELREARAMLRIILMRONGEASVPIEPSQQL 427
DB 61 SCVSILTSEKVLHATEPINEAIKELREARAMLRIILMRONGEASVPIEPSQQL 120
OY 428 DSTMSAEGVLLAGVAGGDFDAIFAITLGDSTKLTQANSSINNALVREPHVCLE 486
DB 121 DATMNEGVLLAGVAGGDFDAIFAITLGDSTKLTQANSSINNALVREPHVCLE 179

RESULT 4
US-10-369-493-3740

; Sequence 3740, Application US/10369493
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.

```

: APPLICANT: Goldman, Barry S.
: APPLICANT: Chen, Xianfeng
: TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
: TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
: FILE REFERENCE: 38-10(52052)B
: CURRENT APPLICATION NUMBER: US/10/369,493
: CURRENT FILING DATE: 2003-02-28
: PRIOR APPLICATION NUMBER: US 60/360,039
: PRIOR FILING DATE: 2002-02-21
: NUMBER OF SEQ ID NOS: 47374
: SEQ ID NO 3740
: LENGTH: 402
: TYPE: PRT
: ORGANISM: Neurospora crassa
: US-10-369-493-3740

Query Match      20.58; Score 529; DB 6; Length 402;
Best Local Similarity 33.88; Pred. No. 1.6e-39;
Matches 172; Conservative 59; Mismatches 134; Indels 144; Gaps 23;

: 6 SAPGKVLMTGGYLVEKPNAGLVSTNARFYAIKPINEEVKPEPMAMKMTDVKLTSPQ.65
: 3 SAPGKVLGAGYIVLDNRITGLVGLSARIVISRDIO-----ASPGVHINEITVRSPOF 57
: 66 SRESM---YKLS-----LNHLTLQ-SVASDSRNPFEVHAIOVAIA-----AAHLATEKD 111
: 58 LKAEMRYGYHLADEBDGKIKITQLOGGASAKGNPFVETTLNVALTYTRRAAGTTO- 115
: 112 KESHLKLLQGLDITILGSDNFYSYRNOIESAGLPITPESIGTLAPASTFPAESNGA 171
: 116 -----TLKPVLTLLIADNDYS-----SPTN-----NSNMAAGKS 146
: 172 NSK-----PEVAKTGLGSSAAMTAVYVALL-HYL--GYVDDSDPCKEKGCSDL 219
: 147 GGSFPAAYSTLEAHKHTGLGSSAALVATLALSHLSDLPDLA--TDEGK----- 198
: 220 DVHIMIAQTSCHLAQGVSGFDVSCAVYGSQRYVRESPEVLSFAQVAVTGLP-----L 273
: 199 RIHLNLAQAAHCAQGVSGFDVAAYGSSHYRRFSPTLS--SLPEACKPGPSAKLF 256
: 274 NEVGTILKGMKDK--RTESLPPLANLFLGEPGSGSSITPSVAGVAKKQMSDPEKAR 331
: 257 SVVNGKAESQMDTEVYKDAVSLPKGVAVRMCDDC--GSQTVGVAKQVLARAAQPKKAK 315
: 332 ENMONLSDANLETKLNDLSKLAKHMDVLYRVIKCSVLTSEKMWLHATEPINEAIIK 391
: 316 ELMELEQARN-----EKLA-----SVLKEGK-----TE----- 338
: 392 ELLEAREAMLRIRILMROMGEASVPIEPESOTOLDS-TMSAGVLLAGVPGAGGPDAT 450
: 339 ---EIRPAVHAIRRELVRKMGESGVPIEPDSQKELLDALEGVGAGVPGAGGYDA- 394
: 451 FAITLGDGSGTULTQAMSHNVALLVRED 479
: 395 -----LALLVRD 402

RESULT 5
US-10-369-493-2162
: Sequence 2162, Application US/10369493
: GENERAL INFORMATION:
: APPLICANT: Cao, Yongwei
: APPLICANT: Hinkle, Gregory J.
: APPLICANT: Slater, Steven C.
: APPLICANT: Goldman, Barry S.
: APPLICANT: Chen, Xianfeng
: TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
: TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
: FILE REFERENCE: 38-10(52052)B
: CURRENT APPLICATION NUMBER: US/10/369,493
: CURRENT FILING DATE: 2003-02-28
: PRIOR APPLICATION NUMBER: US 60/360,039
: PRIOR FILING DATE: 2002-02-21

```

```

: NUMBER OF SEQ ID NOS: 47374
: SEQ ID NO 2162
: LENGTH: 423
: TYPE: PRT
: ORGANISM: Schizosaccharomyces pombe
: US-10-369-493-2162

Query Match      18.88; Score 485; DB 6; Length 423;
Best Local Similarity 30.18; Pred. No. 1.8e-35;
Matches 151; Conservative 83; Mismatches 167; Indels 100; Gaps 21;

: 6 SAPGKVLMTGGYLVEKPNAGLVSTNARFYAIKPINEEVKPEPMAMKMTDVKLTSPQ. 64
: 3 SAPGKVLGAGYIVLDPQSGVIGLFAKGYASTTLLD-----KCGTVRVKSPQF 53
: 65 LSRESMYKLSLN-----HLLOSASDSRNPFVEHAIOVAIAAHLATEKDKESHLK 118
: 54 INAEWLYNIDWTVSPIRVHQIYENCLEKRNPNFVQALFYVI-NYFESTGR-----OPL 107
: 119 LLOGLDITILGSDNFYSYRNOIESAGLPITPESIGTLAP---FASTFPAESNGANSKP 175
: 108 CWODLQVTLQVDMAY-HQPO-----LKPDO--TSYPKENFLNCTIG----- 146
: 176 EVAKTGLGSSAAMTAVYVALLHYLGVDLSDPC--KEGKFGCSDDVYHIMIAQTSCLA 233
: 147 QVHKGTGLGSSAAMTISLIGSL--FLSLRRLTDDTGDKSLKIDDSKYIVHNLQALHCSA 204
: 224 QKVGSGFDVSCAVYGSQRYVRESPEVLSFAQV-----AVTGLPINEVGTILKGMKDKR 289
: 205 QKVGSGFDVGAAMWGSQIYRRFDPKLEOLLVYPDQIKININSTEIRKIVSKMSV- 263
: 290 TEFSLPPLANLFLGEPGSGSSITPSVAGVAKKQMSDPEKARENQNLSDANLETKLN 349
: 264 VPQQLPATYCLMGDV-AGGSSTPGMWKVKQOMKENPEEK---NCFD----- 308
: 350 DLKSLAKHMDVYLRV--IKCSVLTSEKMWLHATEPINEAIIKELLEAREAMLRIRILM 407
: 309 -----DLYSRVLSIKKCFI-----SESDISELQSFRIIRILQRTV-- 347
: 408 ROMGEASVPIEPESQOTOLDSITMSAGVLLAGVAGGAGPAIFAITLGDG--TKLQQA 465
: 348 -----EAKVDIEPKQNTILNDIEQLPGVIGVAGGAGGPAQFLAIINHIEIENVIKT 402
: 466 WSSHNVALLVREDPHVCLE 486
: 403 WKDCGVPMVDYSPAFDGLAVE 423

RESULT 6
US-09-675-784A-13279
: Sequence 13279, Application US/09675784A
: GENERAL INFORMATION:
: APPLICANT: HARE, ROBERTA S.
: APPLICANT: SHAW, KAREN J.
: APPLICANT: SHIMER JR., GEORGE H.
: APPLICANT: KESSLER, MARCO
: APPLICANT: NOLLING, JORR
: APPLICANT: ZENG, QIANDONG
: APPLICANT: GREENE, JONATHAN R.
: TITLE OF INVENTION: ASPERGILLUS FUMLIGATUS NUCLEIC ACIDS AND POLYPEPTIDES,
: TITLE OF INVENTION: AND USES THEREFOR
: FILE REFERENCE: 2976-4020US1
: CURRENT APPLICATION NUMBER: US/09/675,784A
: CURRENT FILING DATE: 2000-09-29
: PRIOR APPLICATION NUMBER: 60/156,338
: PRIOR FILING DATE: 1999-09-29
: NUMBER OF SEQ ID NOS: 13925
: SEQ ID NO 13279
: LENGTH: 491
: TYPE: PRT
: ORGANISM: Aspergillus fumigatus
: US-09-675-784A-13279

```

```

Db 8 SAPKALLAGGYLVLDTRKEAFVYVGLSARMAHVAHPYSGLOGSDKKEFVRNKSQCFKDGEM 67
QY 56 TDVAVLTPSLRSREMYKLSLNHLTLOS-----VSASDSRNFVYHAHQYALAAHNLATKED 111
Db 68 -----LYHISPKSGFLIPVSIIGSKNPEIKVIANVSTF----- 101
QY 112 KESLHKLLOGLDITTLIGSNDPFXSYRNOISAGLPLTPESLGTLPAPASTTFMAAESNGA 171
Db 102 KPNDDYCNRLPYFIDLFSD--AYHSQDS-----VTEHRG-----NRRLSFHS 144
QY 172 NSKPEVAKTGLGSSAAMTTAVVAALLHYLGLVLDSDPCKEGKFGCSLDLVIHMAIOTSHC 231
Db 145 HRIEVEPKTGLGSSAGLVTLTALASFF--VSDLENNVDKX-----EVIHMLAQVANC 197
QY 232 LAQGVSGGDEVSCAVYGSORIVYRFSPEVLSFQAVAVTGLPLNEVIGTILKGR----- 284
Db 198 QAOCKISGDEPVAAAAGSIRYRFRPALIS-----NLP--DIGSATYSGKLAAHLD 247
QY 285 ---DNKRTETSLPLNNLFLGEPGSGSSPSWVGAVKKWQMSDPEKARENQONLSDAN 341
Db 248 EEDNNITTKSNHPLSGLLTMGDI--KNGSETVLYQKVKWYSHMPEISKITTELIDHAN 306
-QY 342 LELETKLNDLSKLAKDHMDVYLRYIKSC--SVLTSEKVVHATEPINALIKELLEAREA 399
Db 307 SRPFDGLSKLDRLEHTHDYSDQIFESLENDCTCKY-----PITEVROA 353
QY 400 MLRIRILMRONGEASVPLEPESOTOLLDSMTSABEYVLAVGPGAGFDAIFAITLGDG 459
Db 354 VATIRSFRRKITKESGADIEPPVQTSLLDCQCLKGLVTLCLIPAGAGYDAIVITRKQVD 413
QY 460 TKLTQAMSSH--NVLALLVREDPHGVCLESGDPRT 492
Db 414 LRAQTANDKRFKQVQWLDVTOADMGVKKE--DPET 447

RESULT 8
US-10-369-493-1909
; Sequence 1909, Application US/10369493
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 1909
; LENGTH: 451
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-10-369-493-1909

Query Match 17.6%; Score 453.5; Db 6; Length 451;
Best Local Similarity 30.1%; Pred. No. 1,5e-32;
Matches 155; Conservative 72; Mismatches 185; Indels 103; Gaps 18;

QY 6 SAPKGLVMTGGYLVLEKPRNAGLVLTNARFYAIYKPKNE-----EYKPESSAMK---W 55
Db 8 SAPKALLAGGYLVLDTRKEAFVYVGLSARMAHVAHPYSGLOGSDKKEFVRNKSQCFKDGEM 67
QY 56 TDVAVLTPSLRSREMYKLSLNHLTLOS-----VSASDSRNFVYHAHQYALAAHNLATKED 111
Db 68 -----LYHISPKSGFLIPVSIIGSKNPEIKVIANVSTF----- 101
QY 112 KESLHKLLOGLDITTLIGSNDPFXSYRNOISAGLPLTPESLGTLPAPASTTFMAAESNGA 171
Db 102 KPNDDYCNRLPYFIDLFSD--AYHSQDS-----VTEHRG-----NRRLSFHS 144
QY 172 NSKPEVAKTGLGSSAAMTTAVVAALLHYLGLVLDSDPCKEGKFGCSLDLVIHMAIOTSHC 231
Db 145 HRIEVEPKTGLGSSAGLVTLTALASFF--VSDLENNVDKX-----EVIHMLAQVANC 197
QY 232 LAQGVSGGDEVSCAVYGSORIVYRFSPEVLSFQAVAVTGLPLNEVIGTILKGR----- 284
Db 198 QAOCKISGDEPVAAAAGSIRYRFRPALIS-----NLP--DIGSATYSGKLAAHLD 247
QY 285 ---DNKRTETSLPLNNLFLGEPGSGSSPSWVGAVKKWQMSDPEKARENQONLSDAN 341
Db 248 EEDNNITTKSNHPLSGLLTMGDI--KNGSETVLYQKVKWYSHMPEISKITTELIDHAN 306
-QY 342 LELETKLNDLSKLAKDHMDVYLRYIKSC--SVLTSEKVVHATEPINALIKELLEAREA 399
Db 307 SRPFDGLSKLDRLEHTHDYSDQIFESLENDCTCKY-----PITEVROA 353
QY 400 MLRIRILMRONGEASVPLEPESOTOLLDSMTSABEYVLAVGPGAGFDAIFAITLGDG 459
Db 354 VATIRSFRRKITKESGADIEPPVQTSLLDCQCLKGLVTLCLIPAGAGYDAIVITRKQVD 413
QY 460 TKLTQAMSSH--NVLALLVREDPHGVCLESGDPRT 492
Db 414 LRAQTANDKRFKQVQWLDVTOADMGVKKE--DPET 447

```

```
OY 172 NSKEPVATGIGSSAAMTTAVVAALHLYGVVLDSDPCKEKGCGSDUDVIMHMQTSC 231
      || ||||| : : : : : || : : : : || : : : : || : : : : ||
Db 145 HRIEVPKGTGSSAGLVTTVITLASF- VSDLENNVDRK-----EVHNLQAVHC 197
OY 232 LAQKVGSGFVSCAVYGSQRYVRSPVLSFAOVAATGLPLNEVIGILGK----- 284
      ||||| : : : : : || : : : : || : : : : || : : : : ||
Db 198 QAQKIGCGFDVAAAAGSIRYRRFPALIS-----NLP---DIGSATYGSKLHLVD 247
OY 285 ---WDMKRTESLPPLMLFLGEGCGSSGSSSTPSWGAIVKQMSPEKARENMQLSDAN 341
      : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 248 EEDNNITKSNHLPSGLLMGDI-KNGSEFVKLVQKKNYDSIMPESLKIYELDHAN 306
OY 342 LELTKLNDLSKLAKDHDVLRVYKSC--SVLTSEKVLHATEPINEAIIKELLEAREA 399
      : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 307 SREWDGSLKDLRHLHETHDYSDQIFESLERNDCTCOKT-----PEITEVDA 353
OY 400 MLRIRILRONGEAAVPIEPESOTQLDSTMSAEGVLLAGVPGAGFDATFAITLDSG 459
      : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 354 VATIRSRFRKTKESGADIEPPVQTSLLDCOTLKGVLTCILPGAGYDAIAVITKQDVD 413
OY 460 TKLQAMSH--NVALLYREDPHGVCLESDDPT 492
Db 414 LRAQTANDRKFSKVQMDLVTAQDWGVREK-DPEY 447

RESULT 9
US-09-909-745A-20
: Sequence 20, Application US/09909745A
: GENERAL INFORMATION:
: APPLICANT: Falco, S. Carl
: APPLICANT: Famodu, Omolayo O.
: TITLE OF INVENTION: Squalene Synthesis Enzymes
: FILE REFERENCE: B1112 US CIP
: CURRENT APPLICATION NUMBER: US/09/909,745A
: CURRENT FILING DATE: 2001-07-20
: PRIOR APPLICATION NUMBER: 09/433,242
: PRIOR FILING DATE: 1999-11-04
: PRIOR APPLICATION NUMBER: 60/107,241
: PRIOR FILING DATE: 1998-11-05
: NUMBER OF SEQ ID NOS: 26
: SOFTWARE: Microsoft Office 97
: SEQ ID NO 20
: LENGTH: 67
: TYPE: PRT
: ORGANISM: Oryza sativa
US-09-909-745A-20

Query Match      8.4%; Score 216; DB 5; Length 67;
Best Local Similarity 60.0%; Pred. No. 2.8e-12;
Matches 39; Conservative 12; Mismatches 14; Indels 0; Gaps 0;

OY 309 GSSTPSWGAIVKQMSPEKARENMQLSDANLETKLNDLSKLAKDHDVLRVYKS 368
      ||||| : : : : : || : : : : || : : : : || : : : : ||
Db 1 GSSTPSWGAIVKQMSPEKARENMQLSDANLETKLNDLSKLAKDHDVLRVYKS 60
OY 369 CSVL 373
      || ||
Db 61 CSRL 65

RESULT 10
US-09-909-745A-16
: Sequence 16, Application US/09909745A
: GENERAL INFORMATION:
: APPLICANT: Falco, S. Carl
: APPLICANT: Famodu, Omolayo O.
: TITLE OF INVENTION: Squalene Synthesis Enzymes
: FILE REFERENCE: B1112 US CIP
: CURRENT APPLICATION NUMBER: US/09/909,745A
: CURRENT FILING DATE: 2001-07-20
: PRIOR APPLICATION NUMBER: 09/433,242
: PRIOR FILING DATE: 1999-11-04
```

```
      PRIOR APPLICATION NUMBER: 60/107,241
      PRIOR FILING DATE: 1998-11-05
      NUMBER OF SEQ ID NOS: 26
      SOFTWARE: Microsoft Office 97
      SEQ ID NO 16
      LENGTH: 54
      TYPE: PRT
      ORGANISM: Zea mays
US-09-909-745A-16

Query Match      7.9%; Score 204; DB 5; Length 54;
Best Local Similarity 74.1%; Pred. No. 2.4e-11;
Matches 40; Conservative 8; Mismatches 6; Indels 0; Gaps 0;

OY 397 REAMRIRILRONGEAAVPIEPESOTQLDSTMSAEGVLLAGVPGAGFDAT 450
      ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
Db 1 RDACLEIRLHMEKMGIAAGVPIEPSQTRLDDATNMMEGVLLAGVPGAGFDAT 54

RESULT 11
US-09-909-745A-18
: Sequence 18, Application US/09909745A
: GENERAL INFORMATION:
: APPLICANT: Falco, S. Carl
: APPLICANT: Famodu, Omolayo O.
: TITLE OF INVENTION: Squalene Synthesis Enzymes
: FILE REFERENCE: B1112 US CIP
: CURRENT APPLICATION NUMBER: US/09/909,745A
: CURRENT FILING DATE: 2001-07-20
: PRIOR APPLICATION NUMBER: 09/433,242
: PRIOR FILING DATE: 1999-11-04
: PRIOR APPLICATION NUMBER: 60/107,241
: PRIOR FILING DATE: 1998-11-05
: NUMBER OF SEQ ID NOS: 26
: SOFTWARE: Microsoft Office 97
: SEQ ID NO 18
: LENGTH: 54
: TYPE: PRT
: ORGANISM: Zea mays
US-09-909-745A-18

Query Match      7.9%; Score 204; DB 5; Length 54;
Best Local Similarity 74.1%; Pred. No. 2.4e-11;
Matches 40; Conservative 8; Mismatches 6; Indels 0; Gaps 0;

OY 397 REAMRIRILRONGEAAVPIEPESOTQLDSTMSAEGVLLAGVPGAGFDAT 450
      ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
Db 1 RDACLEIRLHMEKMGIAAGVPIEPSQTRLDDATNMMEGVLLAGVPGAGFDAT 54

RESULT 12
US-10-282-122A-56929
: Sequence 56929, Application US/10282122A
: GENERAL INFORMATION:
: APPLICANT: Wang, Liangsu
: APPLICANT: Zamudio, Carlos
: APPLICANT: Malone, Cheryl
: APPLICANT: Haselbeck, Robert
: APPLICANT: Ohlsen, Karl
: APPLICANT: Zyskind, Judith
: APPLICANT: Wall, Daniel
: APPLICANT: Trawick, John
: APPLICANT: Carr, Grant
: APPLICANT: Yamamoto, Robert
: APPLICANT: Forsyth, R.
: APPLICANT: Xu, H.
: TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
: FILE REFERENCE: EITRA.034A
: CURRENT APPLICATION NUMBER: US/10/282,122A
: CURRENT FILING DATE: 2003-02-20
: PRIOR APPLICATION NUMBER: 60/191,078
: PRIOR FILING DATE: 2000-03-21
: PRIOR APPLICATION NUMBER: 60/206,848
```



```

Db 4 IEVTPCKLFIAGEYAVEGRHPIIAVADQFVTVTEETDEGSIOSAOYSSLPIMRT- 62
QY 63 POLSRESMYKLSLNHLTLQSVASDSNRNPFVEHAIOYAIAAHATEKDESKLKLLOG 122
Db 63 ---RRNELVLDIR-----ENPF-----HYLAAILH-TERYAOEONKEL--- 98
QY 123 LDITILGSDNFYSYRNOIESAGLPLTPRESLGTLLAPFASITFNAAESNGANSKPREVATGL 182
Db 99 -----SFYHLK-----VSELDSSNG-----RKGL 119
QY 183 GSSAAMTAAYVAAL--LHYLGVDLSDPCKEKGKCGSDLDVIMIAOTSICLAOGKYGSG 240
Db 120 GSSGAVTVGTAKLNIYFDLGL-----ENERIFKLSALAHAYVGN--GSC 163
QY 241 FDVSCAYGSGORYRFSPEVLSFAQAVAVTGLPLNEVIGTILKGMKDKRT--EFLPLPLMN 299
Db 164 GDIAASCTGG--WIAFS---TFHDWVNOKVATEITLIDLAMPPELMFPLKVPOLR 217
QY 300 LFLGEPGSGSSTPSPNAGAVAKKQMSDPEKARENMQNLSDANLELETKLANDSKLAKDHW 359
Db 218 LLIGWTS--PASTSDLVDR---HQSKEKQAEQOFLMSRLCVELMINGFN----- 266
QY 360 DYLIRVTKSCSVLTSEKMWLHATEPINEAIKELLEAREAMLRIIRIMONGEAASVPIE 419
Db 267 -----TGKIS-VIOKOITKNRO-----LLAELSLTGAVIE 296
QY 420 PESOTOLDSTMSAEGVLGAVPGAGFDALFATLIGDSG--TKLTQAMSSHNVAL 474
Db 297 TEALKNLCDLAESTYGA--AKSSGAGGDCGIVIFROKSGILPLMTAMEKDGITPL 350

```

RESULT 15

```

PCT-US02-36122-42
; Sequence 42: Application PC/TUS0236122
; GENERAL INFORMATION:
; APPLICANT: Murphy, Ellen and Projan, Stephen, J.
; TITLE OF INVENTION: Allolococcus otitidis Infectious Disease Targets
; FILE REFERENCE: Application 1
; CURRENT APPLICATION NUMBER: PCT/US02/36122
; CURRENT FILING DATE: 2002-11-25
; NUMBER OF SEQ ID NOS: 106
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 42
; LENGTH: 362
; TYPE: PRT
; ORGANISM: Allolococcus otitidis
PCT-US02-36122-42

```

Query Match 6.0%; Score 154; DB 1; Length 362;

Best local Similarity 20.8%; Pred. No. 2,2e-05;

Matches 103; Conservative 67; Mismatches 168; Indels 158; Gaps 20;

```

QY 8 PGKVLMTGGLVLEKPNAGLVLTNARFYAI---VKPINEVKP-----ESMAKMTDV 58
Db 8 PGKLYLAGEYAVVTPYAGAILL-TVSRVLTLDIWTETSPDQASVRSQTYGNOAYAMERLD- 65
QY 59 KLTSPQLSRESMYKLSLNHLTLQSVASDSNRNPFVEHAIOYAIAAHATEKDESKLHL 118
Db 66 -----GIFSRKDWHP--HLVE----- 81
QY 119 LLOGLDITILGSDNFYSYRNOIESAGLPLTPRESLGTLLAPFASITFNAAESNGANSKPREV- 177
Db 82 -----TVIQTYEAY-----TESLSLPL-----KSYGIQIKSQLD 110
QY 178 ---AKTGLSSAAMTAAYV--AALLHYLGVDLSDPCKEKGKCGSDLDVIMIAOTSHCL 232
Db 111 YQKKIGIGSSGAVTIIVIRGLSLYDLHLKLDID-----IFKLAIAAHIQ 155
QY 233 ACGKVGSGFDVSCAVYGSORVYRFSPEVLSFAQAVAVTGLPLNEVIGTILKGMKDKRT- 291
Db 156 LKSKGSGFDLAACCTYTGVIKQSLDREMLQ--EQIS-----NHSIKDLAMPWSLGLDR 208
QY 292 FSLPPLNLFLGEPGSGSSTPSPNAGAVAKKQMSDPEKARENMQNLSDANLELETKLANDL 351

```

```

Db 209 LSLPHDLRLIGWGTGQ--PASTEKLVQAVYR-----OKITRTPIDFQSFILDO- 253
QY 352 SKLAKDHDVYLRVTKSCSVLTSEKMWLHATEPINEAIKELLEAREAMLRIIRIMONG 411
Db 254 ---SOECVDGLVESLSQADSOASLAMI-----OKNRFLKAMG 288
QY 412 EAASVPIEPESOTOLDSTMSAEGVLGAVPGAGFD--AIFATLIGDSGFKLTQAMSSHN 470
Db 289 QSRGVITETKALTYLCLDIYAKYGG--QAKSSGAGGDCGIGLITRESPTEALYREMDAG 346
QY 471 VLLALVREDPRGVCLE 486
Db 347 ILPLRLDIVENGACVD 362

```

Search completed: April 26, 2003, 13:01:55

Job time : 84 secs

THIS PAGE BLANK (USPTO)